

出版社资源的优化配置

摘要

本文研究了出版社资源配置这一决策优化问题。

需要解决的是总社如何将总量一定的书号合理地分配到各个分社，然后再细分到各个分社下属的各门课程中，使得出版社的经济利益达到最大的问题。考虑到不同的优化方案，我们建立了三个模型：

模型一 分阶段优化资源配置：本模型是以增加强势产品的支持力度为原则，对有限的资源进行优化分配。具体分两步完成，先以 9 个学科的强势程度为依据，得到 9 个分社的分配量；然后类似的从 9 个分社到 72 门课程进行分配。用到的方法有：照顾强势产品法、均值调整和置信度法。最终得到的分配方案见附表一，A 出版社的销售额为 2.2×10^7 。

模型二 单目标规划模型：以出版社的经济利益最大化为目标，直接规划 72 门课程的书号数。利用软件 Lingo 求得最终的分配方案见附表一，A 出版社的销售额为 2.5×10^7 。

模型三 统计回归模型：本模型是基于统计回归的方法对书号进行配置。通过回归分析得到了分配书号与申请书号、销售量、市场占有率和满意度之间的关系。得到和上面模型类似的结果，效果还是比较满意的。

并且在各个模型后面，分别对它们的优缺点进行了讨论。

在最后的灵敏度分析中，我们讨论了总社的书号数在一定范围内变化时，它对分配策略及最终出版社的经济利益的影响。

关键字： 书号 强势产品 人力资源 可靠度 统计回归

一、问题重述

出版社的资源主要包括人力资源、生产资源、资金和管理资源等，它们都捆绑在书号上，经过各个部门的运作，形成成本（策划成本、编辑成本、生产成本、库存成本、销售成本、财务与管理成本等）和利润。

某个以教材类出版物为主的出版社，总社领导每年需要针对分社提交的生产计划申请书、人力资源情况以及市场信息分析，将总量一定的书号数合理地分配给各个分社，使出版的教材产生最好的经济效益。事实上，由于各个分社提交的需求书号总量远大于总社的书号总量，因此总社一般以增加强势产品支持力度的原则优化资源配置。资源配置完成后，各个分社（分社以学科划分）根据分配到的书号数量，再重新对学科所属每个课程作出出版计划，付诸实施。

资源配置是总社每年进行的重要决策，直接关系到出版社的当年经济效益和长远发展战略。由于市场信息（主要是需求与竞争力）通常是不完全的，企业自身的数据收集和积累也不足，这种情况下的决策问题在我国企业中是普遍存在的。

本题附录中给出了该出版社所掌握的一些数据资料，请你们根据这些数据资料，利用数学建模的方法，在信息不足的条件下，提出以量化分析为基础的资源（书号）配置方法，给出一个明确的分配方案，向出版社提供有益的建议。

二、问题分析：

本文需通过分析各分社提交的生产计划申请书、其人力资源情况及市场信息，将总量一定的书号分配给各分社，再从分社分配到其下属的各个课程中。需要给出一个明确的分配方案——即各课程所分得书号数，达到使出版社得到最好经济利益的目标。

考虑对市场信息的把握。首先，作为教材类出版物为主的出版社，在追求经济利益尽量大的同时，须考虑市场不同种类教科书的涉猎，以扩大知名度与市场占有率。其次，因各个分社申请时过分夸大需求，按我们的理解是总社因此无法估算出其未来盈利。为回避单一靠申请书号数目确定经济利益的过大不确定性，总社以增加强势产品支持力度的原则来优化资源配置。这里对该原则的理解就涉及到了解题的关键。

因此我们必须给出强势产品的定义，设立判别标准，然后根据已定出的强势产品的排名，以一定的衡量准则对书号进行优化配置。

在具体分配书号时，需考虑一定学科覆盖率为约束，所以我们要解决的问题归结为：给出一个明确的分配方案，使各分社分得的书号数在满足上、下限、人力资源的约束下，使得A出版社获得的经济效益最大。在计算出版社最终的利润或者营业额时，只能根据已经决策出来的书号分配方案来算，所以找到一个合理的书号与销售量之间的关系也应该是本题的一个重点。

三、符号约定与模型假设：

3.1 符号约定：

X_{ij} ：2006年第*i*个分社第*j*门课程申请的书号数。

x_{ij} ：2006年第*i*个分社第*j*门课程实际获得的书号数。

H : A 出版社拥有的书号数。

X_i : 2006 年第 i 个分社分得的总书号数。

n_i : 第 i 个分社所包含的课程数, $i=1\ldots 9$ 。

a_{ij} : 第 i 个分社第 j 门课程的价格均值。

m_{ij} : 2006 年第 i 个分社第 j 门课程对应的单位书号的销售量。

x_{ijk} : 第 k 年第 i 个分社第 j 门课程实际获得的书号数。

m_{ijk} : 第 k 年第 i 个分社第 j 门课程对应的实际销售量。

\bar{m}_{ijk} : 第 k 年第 i 个分社第 j 门课程单位书号对应的平均实际销售量。

\tilde{m}_{ijk} : 第 k 年第 i 个分社第 j 门课程对应的计划销售量。

h : 所有教材的利润率 (常数)。

Q : A 出版社所得总销售额。

Q_i : 第 i 个分社的销售额。

P : A 出版社所获得的总利润。

P_i : 第 i 个分社的利润。

S_i : 第 i 个分社最多能提供的人力资源。

其它符号见文中相关地方。

3. 2 模型假设:

1. 不考虑人力资源在分社间的调动, 认为各部门的人员数固定。
2. 调查表中的数据作为典型数据, 可以准确反映整个市场的情况。
3. 出版社定价时保持所有教材利润率相同, 在此原则上制定教材单价。
4. 总社在分配书号时, 至少保证分给各分社申请数量的一半; 分社在分配给各个课程时, 也至少分给各课程申请量的一半。
5. 某年的某门课程单位书号销售量是基本不变的。

四、模型的建立与求解:

4. 1 模型准备:

1. 人力资源约束的确定: 各分社的人力资源由策划人员、编辑人员、校对人员中总资

源最小的一个决定。则第 i 个分社最多能提供的人力资源为： $S_i = \min\{A_i, a_i,$

$B_i, b_i, C_i, c_i\}$ 。其中的 $A_i, B_i, C_i, a_i, b_i, c_i$ 分别表示第 i 个分社各类人力资源数和工作能力。

2. 总社书号总数由前几年的分配方案可以算出： $A=500$

4.2 模型的分析与建立

总社在实际分配资源的时候，由于各个分社出于本社利益或其他原因考虑，分社会主观夸大申请的书号数，提交的需求书号量远大于总社的书号总量。因此，总社一般以增加强势产品支持力度的原则优化资源配置。资源配置完成后，各个分社才根据分配到的书号数量，再重新对学科所属每个课程作出出版计划，付诸实施。

由此可知，我们可以将分配书号的问题分两步考虑。第一步是总社分配书号给分社，第二步是分社把得到的书号分给各门课程。建立了如下的模型。

4.2.1 模型一：分阶段优化资源的配置

我们分两个阶段对资源的配置进行优化：

第一阶段的优化：首先确定 9 类学科的强势程度，然后根据增加强势产品支持力度的原则进行优化。

由于出版社最终追求的是经济利益最大，因此，各学科利润 P_i 是决定强势产品的首要因素。由 $P_i = \eta Q_i$ 可知，最大的利润 P_i 等价于最大的销售额 Q_i 。因此我们认为，销售额大小的顺序即是我们需要确定的强势程度的顺序。我们根据附表 4 中的各课程均价及附表 3 中各课程销售量，可以求得各门课程在 5 年内的销售额，再将同一类学科的不同课程相加，即可以得到以下各类学科在五年内的销售额：

	2001	2002	2003	2004	2005
数学类	4962621	5141475	6066546	7869026	9816685
两课类	1735207	2003824	2449023	3142197	3814372
经管类	550254.5	574153.7	765089.7	983145.1	1154493
英语类	436109.9	479209.7	580882.3	752415.9	982840
机械能源类	261885	302400.2	374716.2	466955.7	632717
计算机类	254856.4	266532.6	345299.5	434312.9	533099
地理地质类	172589.7	198806.1	240672.5	287983.1	380361
环境类	145399.8	167561.5	194769.7	264670	372728
化学化工类	88337.27	96252.86	122967.1	159501.3	202736

从上表可以看出，在五年内无一例外地，各学科类的销售额都有一个明确的大小关系即：

数学>两课>经管>英语>机械能源>计算机>地理地质>环境>化学化工 (a)

在第一阶段的优化我们考虑如下的方法：

法一 照顾强势产品法

首先我们简单地将上述关系看作是各学科强弱程度的关系，以下面的原则对总社的 500 个书号进行分配：

1. 为保持工作连续性和对各分社计划一定程度上的认可，A 出版社在分配书号时至少保证分给各分社申请数量的一半。所以，首先为每个分社分配其申请书号数的一半。（这是每个分社分配的下限）
2. 再将剩下的书号按照强势产品进行分配：必须最大限度地满足强势产品的上限。其上限 $M = \min(\text{申请量}, \min(\text{人力资源}))$ 。所以，将剩下的书号根据上面的强弱程度，

从强到弱，依次使分配量达到它的上限。直到 500 号全部分完为止。
这样得到如下的分配方式：

数学	两课	经管	英语	机械 能源	计算 机	地理 地质	环境	化学 化工
120	72	56	102	40	55	20	20	15

显然，这种估优化方法是相当不准确的，因为首先它只根据一种因素来决定强势产品，而且优化的时候都是简单地达到其上限，而排在后面的学科并没有得到优化，只是保持了其初始的下限值。

法二 均值调整法

我们观察前五年实际分配的书号数可以看到，五年内分配给各个分社的书号数基本上趋于平稳，由此我们联想到均值调整法。具体的做法是：第一步首先算出各类学科五年内分得书号的均值，每个分社分得的书号为这个平均值；再根据强势产品的排名作如下的调整：

1. 如果分得的书号数大于其上限，则分得的书号数取为其上限；
2. 如果分得的书号数小于其下限，则分得的书号数取为其下限；
3. 如果最后书号总和大于本分社所分得的书号总数，则从上面 a 式最后一个开始，减少它的书号数，直到其下限；
4. 如果最后书号有剩余，则从上面 a 式最前面一个开始增加，直到其最上限。

这样得到如下的分配方案：

数学	两课	经管	英语	机械 能源	计算 机类	地理 地质	环境	化学 化工
120	72	43	85	40	70	24	25	21

这种分配方案和前一种类似，都是极端地只满足了强势学科类的最大值，而排在后面的学科都没有得到优化。但是这种方法对后面的学科类的书号分配照顾得比较好，因为在给定初始值时是用的每类学科的均值。

法三 置信度法

前面两种方法失败的原因都是因为：在找出了强势产品的排名后，都只是极端地考虑了分配给每个分社的书号上限就是申请量和人力资源的最小值。为了排除这种极端解，重新确定每个分社分配到的书号的上限值，我们引入了置信度的概念。

$$\text{置信度 } S = \frac{\text{实际销售量}}{\text{计划销售量}}$$

它反映了分社在定制计划时的准确度， S 越大说明分社的计划越准确，其申请的书号数的可信度越高， S 就代表了它的可信程度。如果是强势产品，总社在分配时，就应该尽量达到可信度范围内的最大值；反之亦然。可信度的上限就等于 S 乘以申请书号数 X_i ，所以在以增加强势产品支持力度的原则进行分配时，预先分配书号数的上界即是

$$M_i = \min(S \times X_i, P_i)。$$

根据以上的分析，我们需要算出各类学科在前五年内的平均置信度，

$$\text{平均置信度} \bar{S} = \frac{\sum_{i=1}^5 \frac{\text{实际销售量}_i}{\text{计划销售量}_i}}{5} \quad i=1\dots 5，$$

结果如下表所示：

数学类	两课类	经管类	英语类	机械能源类	计算机类	地理地质类	环境类	化学化工类
0.6859	0.6821	0.7032	0.7165	0.7158	0.7249	0.7091	0.6675	0.6923

在改变了分配书号的上限后，我们再使用照顾强势产品法或者均值调整法，对总社的书号进行分配。

第一步得到如下的结果：

数学类	两课类	经管类	英语类	机械能源类	计算机类	地理地质类	环境类	化学化工类	总和
120	49	39	86	43	80	28	27	21	493

从上面的结果以及调整的过程可以看出：

1. 无论用哪种方法给定初值，最后调整后的结果都一样。
2. 对每类学科都有效地照顾到了，因此，对于这种方法，我们可以事先不用排出产品的强势程度。只要设定每类的初始书号为其下限或者均值即可。
3. 最后各类学科分到的书号都达到了它的上限，但是总和却是 493 小于 500 个书号。

由于总社每年都要分出 500 个书号，因此，我们有必要重新调整分配方案，使总社最终能分出 500 个书号。最后剩下书号的原因是分配给数学的书号比往年少许多。

通过观察可以发现，5 年内数学类分得的平均书号为 148.8，而数学类的人力资源分别为：120、144、144。都小于 5 年内它分得的书号数，这在实际中是可以解释的。因为数学知识内容在每年几乎没有更新，而新的书号很多都是用于以前教材的再版，再版时所需要的人力资源明显比出新书所需要的少，特别是策划人员的数目。而人力资源的表格中给出的每个人的工作能力都是指书号用于出版新书的工作能力，所以最后数学分到的书号大于其人力资源。所以对于数学类，人力资源不能作为它的紧约束。

我们设定数学类的书号上限为：144，而其余学科类的书号上限仍然是 M_i ，再重新对书号进行分配。

第二步得到的结果为：

从照顾强势产品法改进得到

数学类	两课类	经管类	英语类	机械	计算	地理	环境类	化学
-----	-----	-----	-----	----	----	----	-----	----

				能源类	机类	地质类		化工类
144	49	39	86	43	80	24	20	15
从均值调整法改进得到								
数学类	两课类	经管类	英语类	机械能 源类	计算机 类	地理地 质类	环境类	化学化 工类
144	49	39	86	42	70	24	25	21

如果我们用与平均值之间的均方误差来评价各种方法的优越性，显然最后得到的置信度法是最好的，而在置信度法中的从均值改进得到的结果更优，因为它本身就是基于平均值来进行调整的。

第二阶段优化

从上面我们已经得到了第 i 类学科所分得的书号数 X_i ，接下来我们对第 i 类学科所包含的 n_i 门课程分配书号。

在对 72 种课程的书号分配进行优化时，从前面模型一的求解过程可以看出，最后一种方法——置信度法是最优的，因此我们使用这种方法对 9 类学科分社所属每个课程进行书号的分配。得到的结果见附表一。

从表中的结果可以看出，由置信度约束的两种方法在分配 72 种课程时没有太大的差别，是因为我们使用了置信度对它们的上限进行了修正。这样得到的结果与前五年的数据之间基本上是相近的，没有太大的波动。

在使用前面的优化方法将 500 个书号分给 72 个学科以后，我们最后计算按照上面的分配方案，出版社获得的营业利润。由于利润率 h 为常数，所以只需要计算出出版社的营业额 Q 。因为营业额 Q 等于销售量与单价乘积，而销售量等于单位书号的销售量乘以书号数。有如下计算公式：

$$Q = \sum_{i=1}^9 \sum_{j=1}^{n_i} a_{ij} * m_{ij} * x_{ij}$$

其中的单位书号的销售量定义为：某种课程的销售量除以该类课程分得的书号数。从附表 2 中所给数据观察得：第 k 年销售的书实际上都是第 $k-1$ 年及以前出版的，因此得到某个单位书号的销售量为

$$\text{单位书号的销售量 } m = \frac{\text{第 } K \text{ 年此类课程的销售量}}{\text{第 } K-1 \text{ 年此类课程的书号数}}$$

要计算销售量，首先必须知道 2006 年对应的某个单位书号的销售量 m_{ij} 。我们可以根据前五年的销售量和分得的书号数，算出 4 组 m_{ij} ，然后再利用这几组单位书号对应的销售量，预测出我们所需的值。

在此考虑最简单的预测方法：使用 MATLAB 中的曲线拟合工具箱进行预测，其基本思想是由已知的几个点拟合出一条曲线，从曲线上计算出下一个点对应的函数值。我

们对 72 种课程分别进行预测，得到如下的 72 组数据：
计算得到：

课 程 代 号	单 位 书 号 对 应 的 平 均 销 售 量	课 程 代 号	单 位 书 号 对 应 的 平 均 销 售 量	课 程 代 号	单 位 书 号 对 应 的 平 均 销 售 量	课 程 代 号	单 位 书 号 对 应 的 平 均 销 售 量
1	268.6	19	4660	37	87.59	55	1265
2	353.5	20	1211	38	1662	56	1715
3	101.1	21	625.7	39	519.3	57	702.4
4	356.3	22	843.5	40	385	58	178.5
5	125.1	23	9904	41	6760	59	914.2
6	416.3	24	1325	42	6384	60	490.5
7	543.6	25	1671	43	3039	61	297.7
8	81.02	26	3004	44	4965	62	657.8
9	466.5	27	621.1	45	35632	63	736.8
10	571.9	28	645.3	46	3644	64	584.5
11	1181	29	392.2	47	2555	65	172
12	1898	30	2086	48	1610	66	1102
13	456	31	681.5	49	840.9	67	603.4
14	793.3	32	224.1	50	853.1	68	379.9
15	517.4	33	1215	51	439.7	69	648.7
16	1048	34	265	52	1067	70	586.2
17	980	35	566	53	1595	71	393.3
18	414.9	36	216.5	54	1265	72	1286

根据前面对各个课程所分配的书号数，可以由上式可以计算出 A 出版社所收入的营业额为：

当代入的是置信度法中均值调整所得到的书号分配可得到： $Q = 2.26 \times 10^7$

当代入的是置信度法中照顾强势所得到的书号分配可得到： $Q = 2.27 \times 10^7$

由结果可以看出，由两者得到的营业额大概一致，但置信度改进的照顾强势法营业额更高，从这一点上看，置信度中的照顾强势法要更优。由于利润率 η 为常数，所以 A 出版社获得的利润为： $P = \eta \times Q = 2.27\eta \times 10^7$ 。

我们从网上查到了一般出版社的利润率在 10% 左右，因此，可以取 $\eta = 10\%$ ，最终得到利润为： $P = 2.27 \times 10^6$ 。

模型评价：

优点：该模型是先对 9 个学科分配书号，然后再对每一学科中的每门课程分配，最终得到 72 门课程的所分配的书号数。分配过程中照顾到了增加强势产品支持力度这一

原则。求解出的分配方案合理。思路比较清晰，

缺点：需做大量的数据统计，求解过程复杂。而且还有题目中用到的 2006 年的单位书号对应的销售量是预测出来的，可能存在误差，对最后结果造成影响。

我们进一步想，可以直接以 72 门课所分得的学号数为决策变量，以经济利益最大为目标，建立一个整体优化的单目标规划模型。

4.2.2 模型二：

$$\begin{aligned}
 & \max \quad P \\
 & s.t. \\
 & \quad P = hQ \\
 & \quad Q = \sum_{i=1}^9 \sum_{j=1}^{n_i} a_{ij} * m_{ij} * x_{ij} \\
 & \quad \sum_{j=1}^{n_i} x_{ij} \leq P_i \quad i = 1, 2, \dots, 9 \quad (1) \\
 & \quad \sum_{j=1}^{n_i} x_{ij} \geq \frac{1}{2} \sum_{j=1}^{n_i} X_{ij} \quad i = 1, 2, \dots, 9 \quad (2) \\
 & \quad \sum_{i=1}^9 \sum_{j=1}^{n_i} x_{ij} = 500 \quad (3) \\
 & \quad \frac{1}{2} \sum_{j=1}^{n_i} X_{ij} \leq x_{ij} \leq \sum_{j=1}^{n_i} X_{ij} \text{ 且为整数} \quad (4)
 \end{aligned}$$

此模型是以 72 种课程的书号分配为决策变量，以 A 出版社整体利润最大为目标，以书号分配时的规则为约束进行规划。

模型具体解释：目标函数即为销售各门课程所得的利润之和最大。Q 为总销售额，h 为所有教材的利润率。即为销售各门课程所得的利润之和最大，Q 为总销售额，h 为所有教材的利润率。 $x_{ij} * m_{ij}$ 即为第 i 个出版社第 j 门课程的平均销售量， $a_{ij} * x_{ij} * m_{ij}$ 为该课程对应的销售额。

- (1) 式为人力资源约束，即分配给第 i 个分社的总书号数小于该分社的最大人力资源；
- (2) 式为为保持工作连续性和对各分社计划一定程度上的认可，A 出版社分配书号时至少分给各分社申请数量的一半；
- (3) 式为要分配的总书号数为 500 个；
- (4) 式表示分配给第 i 个分社第 j 门课程的书号数大于其申请数的一半。

模型的求解：

此模型在求解时，我们使用了规划软件 Lingo 对其进行优化，优化时我们以 Q（总销售额）最大为目标，最后得出的 72 种课程的分配方案见附表一。

最终得到的营业额为： $Q = 2.54 \times 10^7$ 。

A 出版社最终的利润为： $P = 2.54 \times 10^7$ ，

如果取 $\eta = 10\%$ ，则得到的利润为： $P = 2.54 \times 10^6$ 。

从 Lingo 求解出的结果可以看到，营业额比模型一中的几种先设定初值、再进行调整的方法要优化一些，因为 Lingo 在优化时，是统筹考虑 72 种课程的利润总和。而前面的几种方法都是分步优化，即是先考虑总社分到分社，再考虑分社自己内部分到各种课程。而且，前面的分步优化时，第一步的优化只是人为寻找了一个比较合理的可行解，在决策时并没有对目标进行优化。

模型评价：

优点：该模型直接以 72 门课程所分得的书号数为决策变量，以总利润最大为目标，用 Lingo 可以很快求出最优解，方法简便快捷，具有广泛的实用性。

缺点：该模型主要考虑如何分配使得 2006 年的经济利益最大，各课程分得的书号数和课程均价与该课程每个书号对应的平均销售量的乘积成正比。由于考虑的因素不全，如没有考虑市场占有率等因素的影响。

于是我们认真分析调查表（附表 2），通过对数据地处理，综合考虑市场占有率、申请书号数等建立了如下的统计回归模型。

4.2.3 模型三：统计回归模型

问题分析：

考虑分配书号数的相关因素，我们进行分析后认为可能的直间接因素主要有四点：
1. 该类课程的该年申请分配书号。
2. 该课程强势程度，它按前面的理解包含两点（1）该社该课程教材的市场占有率。即考虑到某课程可能在社内盈利不高，但是在市场上占有的占有率，那么我们为巩固市场份额，还是会以一定比率照顾。（2）该课程在自己社所有课程中占的盈利比率。即考虑社内部分号时的盈利目标。
3. 该课程的所属学科人力约束。
4. 该课程的市场满意度。对于第四点的考虑，我们认为如果出版社发现自己某课程相对满意率较高则会延缓再版，该课程相对其它课可能分配不到再版的新号。

所以，我们期望从数学统计的方法找出决定实际书分配号的相关量影响强度的量化关系。

模型准备：

1. 确定出 72 课程前几年的申请书号数。由于我们假设某特定课程年间销售量/书号是基本不变的。所以，根据当年该课程的实际销售量与计划销售量的比值，即计划准确度，我们可以由简单的比例关系，计算出当年该学科计划申请书号数。
2. （1）统计出该课程实际销售额在我社总销售额的比率。即由均价乘销售量作为 所求销售额进行比较。
（2）统计出该社出版该课程教材的市场占有率。
3. 因为我们寻找关系的数据除申请书号数外均为已得到的实际值的统计。虽然它实际得到的书号可能受申请年人力的影响，但是考虑人力数据均值与分配号所需接近，即人力年变化不大，所以我们忽略当时人力约束对分配号的影响。
4. 对于市场满意度的考虑，我们计算出该社该课程教材的年平均满意率。即其年平均满意度除以所有该课程的平均满意度。再于我社内部做归一化处理。相当于考虑各社间与各分社间的满意程度差距。

模型建立：

考虑各相关量单位不一致的情况，我们采用统一的归一化占有率，具体建立如下多元统计回归预测方程：

$$y = f(jh_i^j, sc_i^j, yl_i^j, sa_i^j) \quad (i, 1..72; j, 1..5)$$

其中 jh_i^j 表示第 j 年第 i 课程的计划申请书号数占该年各分社计划申请总数比例。

sc_i^j 表示目标出版社第 j 年第 i 课程各教材总数所占该课程教材总数的占有率。同理，

yl_i^j 表示第 j 年第 i 课程占目标出版社市场销售额的比例。 sa_i^j 表示目标出版社第 j 年第 i 课程教材平均满意率。

模型求解：

实际处理数据时我们发现以单个课程分配书号为决策变量则需统计提取信息过于庞大，且分配号数与于各项可能因素的具体相关程度并不清楚，如果盲目做，会得到一个比较满意的精确的量化关系。然而若得到目标与某因素关系不大，那么之前工作显得十分低效。另外，考虑到分配实际为先分社得到号，再分配给所属课程。所以，我们把以上统计改为按九个学科类别。以同样的思路进行建模求解。

(1) 简化方程如下：

$$\tilde{y} = f(\tilde{j}h_i, \tilde{s}c_i, \tilde{y}l_i, \tilde{s}a_i) \quad (i, 1..9)$$

其中 $\tilde{j}h_i, \tilde{s}c_i, \tilde{y}l_i, \tilde{s}a_i$ 分别表示第 i 类学科的相应量，即上述模型统计主体改为学科即可。

(2) 统计出对应量，对每列做求均值归一化处理。结果为四个 9×5 的向量矩阵。

(3) 考虑数据波动上下较均匀，为再次简化，我们合并数据，求出5列均值，得到 \tilde{y} ，

$\tilde{j}h_i, \tilde{s}c_i, \tilde{y}l_i, \tilde{s}a_i$ 构成 9×4 向量矩阵。

(4) 即 $y = \beta_0 + \beta_1 x_1 + \beta_2 x_2 + \beta_3 x_3 + \beta_4 x_4$ [1] 因变量 y 与自变量 x_1, x_2, \dots, x_4 之间是否存在如

[1] 所示线性关系是需要检验的，显然，如果所有 $|\hat{\beta}_j| (j=1, \dots, 3)$ 都很小，

y 与 x_1, \dots, x_4 的线性关系就不明显，所以可令原假设为：

$$H_0: \beta_j = 0 (j=1, \dots, m) \quad (2)$$

当 H_0 成立时，回归平方和 U 和残差平方和 Q 满足：

$$F = \frac{U/m}{Q/(n-m-1)} \sim F(m, n-m-1) \quad \square \quad F(m, n-m-1) \text{ 在显著水平 } \alpha \text{ 下有 } 1-\alpha \text{ 分位数 } F_{1-\alpha}(m, n-m-1), \text{ 若}$$

$F < F_{1-\alpha}(m, n-m-1)$ ，接受 H_0 ；否则，拒绝。即当 F 值大于一定的临界值时，拒绝原假设，

认为因变量与自变量之间是相关的。

在确定因变量 y 与自变量 x_1, x_2, \dots, x_4 之间存在线性关系后，我们采用逐步回归法求取回归系数，进行回归模型的建立。

首先，对偏相关系数最大的变量做回归系数显著性检验，以决定该变量是否进入回归方程；然后，对方程中的每个变量作为最后选入方程变量的变量求出偏 F 值，对偏 F 值最小的那个变量作偏 F 检验，决定它是否留在回归方程中。重复此过程，直至没有变量被引入，也没有变量可剔除时为止。

用相关性检验发现相关性十分显著，即 $R\text{-square}=0.998634$ ， $F=2192.8$ ， $P=0.0000$ 。

经结果分析发现 $\tilde{s}c_i$ ， $\tilde{s}a_i$ 对方程影响比重很小，所以经回归诊断检验后，我们剔除了此两项。（回归过程残差图与数据见附表）。

得到最终回归方程为

$$y=1.0350*\tilde{j}h-0.0334*\tilde{y}l;$$

代入06年各科计划申请书号数，以及由前五年多项式拟合得到的各科盈利比率。我们得到一组各分科书号分配预测值如下：

计算机类	经管类	数学类	外语类	两课类	机械、能源类	化学、化工类
75	38	144	82	46	41	20
地理、地质类	环境类					
27	27					

模型检验及结果分析：

考虑到对模型准确度的检验，我们用同样的访法采用01—04的数据对05的分配书号数进行了预测。发现由该数据得到的回归方程与05年各科实际分配值误差仅为0.25%。

可见，模型准确度是比较高的。然而，对于统计结果只与当年申请书号数、该学科在目标出版社的盈利率有关的结果还是令我们惊讶。结果表明我们先前考虑到的分配书号时出版社对市场占有率的支持未体现。后来我们分析认为，可能出版社在某学科占有较大规模的市场时，会形成稳定地位而不再担心它社超出。而在市场占有率较小时，可能会考虑增加投资力度，包括增新号数等。因而正负相关关系难以确定。可能存在更复杂的非线性关系。这也是此模型需改进深究的地方。

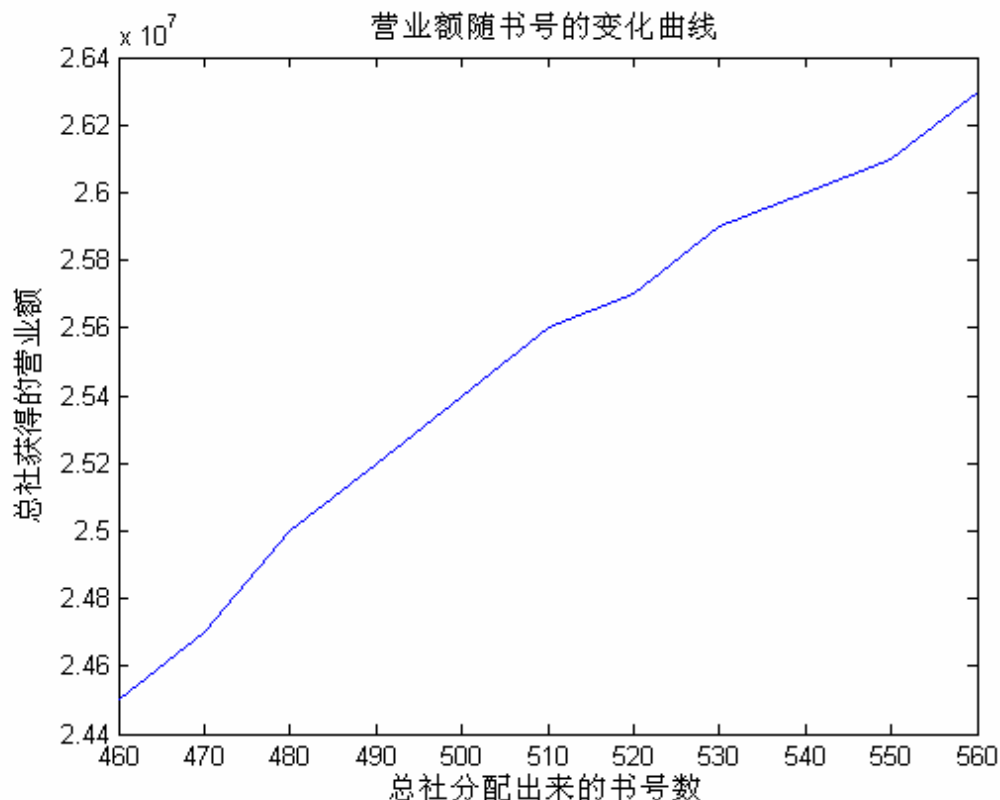
另外，在相对市场面的扩展上出版社考虑其影响远大于盈利比率。这也一定程度上表明教材类出版社的特点，即兼顾稳定学科供应量，市场覆盖面的前提下才尽可能增加盈利率。

五、灵敏度分析：

我们将 A 出版社拥有的书号总数作为灵敏度的研究对象,因为它的数量直接影响到总社为分社分配书号的决策,而且它也是我们将要优化目标的最重要的决定因素。

1. 出版社的营业额与总书号数的关系

我们用 Lingo 程序找到不同书号总数条件下,出版社营业额的变化情况。做出如下的曲线图:



从上面的曲线可以看出,总社的营业额与其分配出的书号基本上成正比关系。由于营业额与利润之间差一个常数比例,所以出版社的利润就与总书号数成正比。

在前面的模型中,由营业额计算公式,不同课程的营业额是由不同的单价和单位书号销售量作为加权系数乘以各课程分配到的书号数,所以总的营业额不一定与总书号数成正比。但是当总社的总书号数增加时,只要总社适当分配其拥有的书号数,总可以使它的营业额提高相应的比例,即利润增加一定的比例。

这就是在实际中各个出版社都想申请到尽量多的书号数的原因。因此,新闻出版局或政府相关部门在分配各个出版社的书号时,要先考查该出版社的综合能力,再分配给它适量的书号数。

2. 分配策略与总书号数的关系

当总书号发生变化时,分配策略并没有太明显的变化规律。简单统计如下:
当书号从 500 降低到 480 再降到 460 这个范围时,分配到的书号数降低的课程号有: 14、31、49、52、62、67,从已知的信息中很难简单地找到他们之间的共同点,所以只知道这些课程对总书号数的降低比较敏感。同样,对总书号数增加比较敏感的课程的代号为: 7、18、35、49、57、63。如果想到得到总书号发生变化时,各种课程的分配方案的改变规律,则需要更多的信息及更深入研究。

参考文献：

- [1] 姜启源、谢金星、叶俊，《数学模型》第三版，北京：高等教育出版社，2003；
- [2] 苏金明、阮沈勇、王永利，《MATLAB 工程数学》，北京：电子工业出版社，2005；
- [3] 晏林，《数学与实验—MATLAB 与 QBASIC 应用》，北京：科学出版社，2005；
- [4] 韩中庚，《数学建模方法及其应用》，北京：高等教育出版社，2005；
- [5] 薛定宇、陈阳泉，《高等应用数学问题的 MATLAB 求解》，北京：清华大学出版社，2004。

附表一：

模型一及模型二的求解结果：72 种课程的具体书号分配方案
置信度法

学科 类别	课程种类	均值 调整	照顾 强势	Lingo 结果
计算机类	C++程序设计	12	13	9
	C 程序设计	11	13	9
	DSP 技术及应用	3	3	2
	Java	3	4	3
	编译原理	3	5	3
	数据结构	11	11	8
	软件工程	9	9	6
	单片机	3	4	3
	多媒体	10	12	8
	人工智能	5	6	4
	总计	70	80	55
经管类	保险	5	6	8
	组织行为学	3	3	4
	证券投资	3	3	2
	西方经济学	3	3	4
	企业管理	4	4	3
	计量经济学	4	4	6
	技术经济学	4	4	6
	财务管理	6	5	4
	管理信息系统	4	4	6
	国际经济学	3	3	4
	总计	39	39	47
数学	离散数学	7	6	6

类	数学分析	25	25	19
	高等数学	35	35	35
	常微分方程	5	4	4
	复变函数	18	18	12
	概率论与数理统计	21	24	17
	近世代数	7	6	6
	经济数学	3	3	3
	微积分	15	15	12
	线性代数	8	8	6
	总计	144	144	120
英语类	大学英语	29	29	40
	法语	3	3	2
	实用翻译教程	1	1	2
	泛读	16	16	11
	计算机英语	5	5	4
	口语	12	12	8
	美国文学	4	4	3
	日语	5	5	6
	商务英语	7	7	5
	语法	4	4	3
	总计	86	86	84
两课类	当代世界经济与政治	3	3	4
	邓小平理论和“三个代表”重要思想	7	7	10
	马克思主义政治经济学原理	5	5	8
	马克思主义哲学原理	6	6	10
	毛泽东思想概论	4	4	6
	思想道德修养	6	6	8
	法律基础	8	8	12
	政治经济学	10	10	14
	总计	49	49	72
机械、能源类	工程制图	12	12	14
	过程控制	3	3	4
	画法几何	7	7	5
	机械设计	11	12	16
	机械原理	3	3	4
	机械制图	6	6	8
	总计	42	43	51
化学、化工类	化学与现代文明	2	2	4
	有机化学	3	2	4
	物理化学	3	2	2
	化工原理	3	2	2

	工程化学	3	2	2
	普通化学	7	5	5
	总计	21	15	19
地 理、 地质 类	城市地理学	4	4	4
	地理信息系统	4	5	8
	地图学	5	4	4
	地质学	5	6	4
	工程地质	3	2	2
	经济地理学	3	3	4
	总计	24	24	26
环 境 类	大气污染控制工程	4	4	8
	水污染控制工程	6	5	5
	环境学	6	4	4
	环境生态学	5	3	3
	环境化学	2	2	2
	环境管理	2	2	4
	总计	25	20	26