

# 基于熵值法的高校分省招生计划编制研究

高雅, 倪建威, 李旭涛, 陈鹏丞, 张艳萍\*

河北工程大学数理科学与工程学院, 河北 邯郸

Email: [zhangyanping@hebeu.edu.cn](mailto:zhangyanping@hebeu.edu.cn)

收稿日期: 2021年4月17日; 录用日期: 2021年5月2日; 发布日期: 2021年5月20日

## 摘要

合理分配高校招生计划对提高高校生源质量具有十分重要的作用与影响。本文利用熵值法对某高校2017~2019年分省招生数据进行分析, 计算出5个指标值的具体权重值: 历年招生计划指标 $I_1$  (0.959925), 招生最低分与控制线差值指标 $I_2$  (0.036203), 调档线与控制线比值指标 $I_3$  (0.000754), 志愿意向指标 $I_4$  (0.003003), 实际招生情况指标 $I_5$  (0.000116), 并以此给出各省份具体招生计划数, 将30个省份分为三大类型, 分别给出分配计划调整意见。利用熵值法对招生计划进行分配, 具有理论依据, 更加客观, 为招生部门提供指导意义更具有说服力。

## 关键词

熵值法, 招生数据, 权重, 招生计划

# Research on the Provincial Formulation of University Enrollment Plan Based on Entropy Method

Ya Gao, Jianwei Ni, Xutao Li, Pengcheng Chen, Yanping Zhang\*

School of Mathematical Science and Engineering, Hebei University of Engineering, Handan Hebei

Email: [zhangyanping@hebeu.edu.cn](mailto:zhangyanping@hebeu.edu.cn)

Received: Apr. 17<sup>th</sup>, 2021; accepted: May 2<sup>nd</sup>, 2021; published: May 20<sup>th</sup>, 2021

## Abstract

Reasonable allocation of college enrollment plan plays a very important role in improving the

\*通讯作者。

quality of college students. In this paper, the entropy method is used to analyze the provincial enrollment data of the University from 2017 to 2019. The weights of the five index values are calculated: enrollment plan over the years  $I_1$  (0.959925), the difference between the minimum score of enrollment and the control line  $I_2$  (0.036203), ratio of shift line to control line  $I_3$  (0.000754), voluntary intention  $I_4$  (0.003003), actual enrollment  $I_5$  (0.000116). According to these five indicators and weights, we give specific enrollment plans for each province. In addition, we divide 30 provinces into three types and give suggestions on adjustment of allocation plans for each type.

## Keywords

Entropy Method, Enrollment Data, Weight, Enrollment Plan

Copyright © 2021 by author(s) and Hans Publishers Inc.

This work is licensed under the Creative Commons Attribution International License (CC BY 4.0).

<http://creativecommons.org/licenses/by/4.0/>



Open Access

## 1. 引言

高校招生计划是国家落实教育规划、调节高等教育规模发展节奏、提高高等教育质量的重要手段，高校招生计划的编制不仅对高校人才培养具有基础性作用，对基础教育也具有导向性作用[1]。高校招生计划编制管理工作决定了学校生源的地域结构，对生源质量具有直接的影响作用，同时也决定了不同省份人才跨区域流动的结构和数量，关系到人才选拔战略[2]。因此，系统全面的分析编制招生计划的影响因素，并采取有效的模型编制招生计划，对于推动生源质量提高，促进高校与区域经济、产业升级的协调发展具有重要的作用与意义。

招生计划的编制是一个复杂的系统工程，涉及生源、地域、政府、高校和社会等多方面因素，要进行系统全面的招生计划编制工作需要众多因素进行全方位考虑，但许多因素难以量化和把握，这就为招生计划的编制工作带来了一定的挑战。近年来，许多专家与学者在高校招生计划编制上开展了丰富的探讨与研究，数学建模的方法早已成为高校招生计划领域研究的亮点[3]。杨卫平[4]等人从生源、高校、政府三个维度出发，运用模糊数学方法进行高校招生来源计划编制。郑庆华[5]等人，从教育行政主管部门和高校两个主体出发，运用二次分配模型，进行普通高校分省招生计划的编制，充分考虑了招生计划编制过程中，决策目标复杂、参与主体较多的特点。陈琳[6]等人从高校自身出发，综合考虑专业办学条件、近三年报到率、报考率、就业率及专业内涵建设情况，建立学校内部各专业招生计划二次分配的测算模型。付辉[1]运用规范分析法对高校招生过程、影响因素及招生数据进行分析，采用灰色预测模型等方法对招生总计划进行预测，采用层次分析法对高校历史招生数进行分析，从而对高校分省分专业招生计划的制订提供指导意见。魏针[7]等人主要从招生计划和办学条件两个方面出发，建立基于破产理论的省属高校招生计划分配模型，但应用到具体某一高校的分配时，需要为专业评估得分，标准因各高校实际情况而异，难度较大。

在目前的招生计划分配研究中，系统全面地建立招生计划编制体制难度依然较大，不同高校受地域、办学条件等因素的影响，招生计划制订也不尽相同。招生计划分配方案一方面基于层次分析法，模糊数学等方法，评价指标难获得，另一方面多采用专家打分，受主观性影响。因此，本文提出利用熵值法对历年数据进行分析，减少主观因素影响。从某高校招生部门出发，结合学校自身特点和实际条件，基于历史招生数据，建立适合本校的分省招生计划编制模型，为学校合理制定招生计划，提高生源质量提供参考。

## 2. 研究方法和模型

### 2.1. 数据来源与指标选取

选取某高校 2017~2019 年招生数据为原始数据,包括 2017~2019 年各省份计划录取数、实际录取数、志愿录取数、录取最低分、控制线、调档线和 2017~2019 年某高校总计划招生数。通过对原始数据指标进行转换得到以下 5 个指标:

**Table 1.** Description of the five index values

**表 1.** 5 个指标值说明

指标	含义	说明
$I_1$	各省份招生计划与学校招生总计划的三年平均值	反映历年各省份招生计划比重情况,是招生分配计划中重要指标
$I_2$	各省录取最低分减各省控制线的三年平均值	分数高低代表生源质量情况,各省份录取最低分与控制线的差值反映所招学生在各省份的位次
$I_3$	各省调档线比各省控制线的三年平均值	调档线高低反映学校位次情况及学生位次情况,与各省省控线比较,使各省份之间可以进行比较
$I_4$	各省份志愿录取数比计划录取数的三年平均值	这一指标代表各省份有意愿报考学生比例情况,我省招生投档遵循志愿优先
$I_5$	各省份实际招生数比计划录取数的三年平均值	这一指标直观的反映出各省份近三年招生录取情况,及招生计划完成情况,侧面反映了招生计划是否合理

在表 1 的五个指标介绍中,其中指标  $I_1$  代表三年各省份平均录取情况,指标值越大,说明该省份在 2017~2019 年录取人数越多,历年招生计划分配情况是未来进行招生计划分配决策的基础;指标  $I_2$  和  $I_3$  从招生分数方面反映各省份 2017~2019 年招生质量情况,指标值越大,说明考生在各省相对排名更靠前,招生质量越好,在招生计划分配中,在历年招生计划分配基础上给予一定权衡;指标  $I_4$  和  $I_5$  从招生数量方面反映各省份 2017~2019 年考生报考意愿情况,指标值越大说明该省份考生报考本校意愿更强,本校在该省份的录取情况越好,计划完成率越高,招生计划分配应充分考虑历年招生完成情况及考生意愿,考生意愿越强烈招生质量越高。本文选取的 5 个指标皆来自于招生部门原始数据,均为正向指标,以历年招生计划分配为基础,考虑招生质量和数量对招生计划分配影响,以理论为依据对各省份招生计划分配进行规范化,使分配更加合理,具有说服力。

### 2.2. 熵值法

熵值法是一种客观赋权方法,根据各项指标观测值所提供信息的大小确定各项指标权重。信息论中,熵是对不确定性的一种度量,信息量越大,不确定性越小,熵越小,反之信息量越小,不确定性越大,熵越大。根据熵的特性,通过计算熵值可以判断一个方案的随机性和无序程度或某个指标的离散程度,指标的离散程度越大,则该指标包含信息量越大,对综合评价的影响越大。因此,可根据各项指标的变异程度,基于信息熵计算各指标的权重,为多指标综合评价提供依据。

### 2.3. 基于熵值法的指标权重计算

利用熵值法对本文所构造的 5 个指标进行权重计算:

1) 构造判断矩阵

设有  $m$  个评价指标对  $n$  个评估对象进行评估,则判断矩阵  $A$ :

$$A = \begin{pmatrix} x_{11} & \cdots & x_{1m} \\ \vdots & \ddots & \vdots \\ x_{n1} & \cdots & x_{nm} \end{pmatrix}$$

其中,  $x_{ij} (i=1, \dots, n; j=1, \dots, m)$  表示第  $i$  个方案中第  $j$  个指标的数值, 本文中有 30 个省份, 5 个指标, 即  $m=5, n=30$ 。

由于各指标含义不同, 计量单位各异, 因而需要对数据进行标准化处理, 消除不同量纲对结果的影响, 对  $x_{ij}$  进行标准化处理:

$$p_{ij} = \frac{x_{ij}}{\sum_{i=1}^n x_{ij}}$$

其中,  $p_{ij}$  为  $x_{ij}$  标准化后的指标数据, 设标准化后的矩阵为  $P$ 。

2) 定义第  $j$  个指标的熵值:

$$e_j = -\frac{1}{\ln n} \sum_{i=1}^n h_{ij} \ln h_{ij}$$

其中,  $e_j$  为第  $j$  项指标的熵值,  $h_{ij} = 1 + p_{ij} / \sum_{i=1}^n (1 + p_{ij})$  表示第  $i$  个指标对第  $j$  个评估对象的特征比重值。

3) 计算第  $j$  项指标的差异系数

对于第  $j$  项指标, 指标值  $p_{ij}$  的差异越大, 对方案评价的作用越大, 熵值就越小, 因此计算第  $j$  项指标的差异系数:

$$g_j = 1 - e_j$$

差异系数越大, 指标越重要。

4) 求权数

设  $w_j$  为第  $j$  项指标的权重, 则:

$$w_j = \frac{g_j}{\sum_{j=1}^m g_j}$$

其中,  $j=1, \dots, m$ , 本文利用 R 软件对 5 个指标的权值进行计算, 结果如表 2:

**Table 2.** The calculation result of index weight

**表 2.** 指标权重计算结果

指标	I <sub>1</sub>	I <sub>2</sub>	I <sub>3</sub>	I <sub>4</sub>	I <sub>5</sub>
权重 $w_j$	0.959925	0.036203	0.000754	0.003003	0.000116

通过表 2 各指标的权重大小可以发现, 指标 I<sub>1</sub> 权重明显高于其它指标, I<sub>1</sub> 代表历年各省招生计划分配比例情况, 通过历年招生数据规律对未来各省招生名额进行分配, 显然历年各省招生比例是最为重要的评价指标, 熵值法的计算结果与实际情况相一致, 表示利用熵值法对各指标进行权重赋值是有效合理的。其它 4 个指标的权重都相对较小, 表明在这 4 个指标下, 各省份的差异较小, 对省份区分并不明显, 但有一定作用。

## 2.4. 分省招生计划人数计算

利用综合指数法思想, 对各省份指标权重累乘、累加, 计算各省份总权重:

$$W_i = \sum_{j=1}^m w_j p_{ij}$$

其中  $W_i$  为第  $i$  个省份的总权重,  $i=1, \dots, n$ 。

设某高校 2020 年总招生计划为  $N$ , 则第  $i$  个省份的招生计划名额为:

$$N_i = W_i \times N$$

本文以 5520 人为例, 计算各省份招生名额分配情况及各省权重如表 3 所示。

**Table 3.** The calculation result of enrolment number in different provinces

**表 3.** 不同省份招生人数计算结果

省份	$W_i$	计划招生数	省份	$W_i$	计划招生数
北京	0.002062	11	河南	0.026662	147
天津	0.012856	71	湖北	0.023044	127
河北	0.552256	3049	湖南	0.022868	126
山西	0.035108	194	广东	0.005206	29
内蒙古	0.019737	109	广西	0.007506	42
辽宁	0.018819	104	海南	0.010343	57
吉林	0.030516	169	重庆	0.006575	36
黑龙江	0.019945	110	四川	0.020789	115
上海	0.002249	12	贵州	0.014358	79
江苏	0.016733	92	云南	0.011412	63
浙江	0.011393	63	陕西	0.035582	196
安徽	0.018674	103	甘肃	0.005984	33
福建	0.012394	68	青海	0.002767	15
江西	0.017886	99	宁夏	0.004979	28
山东	0.027282	151	新疆	0.004014	22

根据表 3 可以发现, 各省份招生人数差距较大, 北京招生人数最少仅为 11 人, 这与历年招生计划息息相关, 2017~2019 年本校在北京市招生仅为 4 人, 但每年招生完成情况较好, 因此在此基础上熵值法对北京市招生进行上调, 从理论角度出发给出更合理的招生计划。河北省招生人数占据总招生计划一半以上, 这也与历年招生数据息息相关。高校招生计划分配受教育局调控影响, 而这一指标不易获得, 但通过挖掘历年招生数据, 可以发现包含了教育局调控这一指标对招生计划影响。

## 3. 结论

本文利用熵值法对某高校 2017~2019 年招生数据进行分析, 得到各省份相对权重值及最终招生计划分配情况。进一步, 通过对比历年招生计划及比例分析各省招生情况, 2017~2019 年及预测的 2020 年各省份招生计划及比例如表 4 和图 1 所示。

Table 4. The proportion and plan of enrollment by province in 2017~2020

表 4. 2017~2020 年各省份招生比例及计划

省份	2017 占比/计划		2018 占比/计划		2019 占比/计划		2020 占比/计划	
北京	0.001	4	0.001	4	0.001	4	0.002	11
天津	0.010	60	0.010	60	0.016	85	0.013	71
河北	0.604	3792	0.579	3396	0.541	2935	0.552	3049
山西	0.033	210	0.036	210	0.039	210	0.035	194
内蒙古	0.020	126	0.020	116	0.021	116	0.020	109
辽宁	0.016	100	0.017	100	0.018	100	0.019	104
吉林	0.029	180	0.031	180	0.033	180	0.031	169
黑龙江	0.018	110	0.019	110	0.019	105	0.020	110
上海	0.002	10	0.002	10	0.002	10	0.002	12
江苏	0.018	110	0.016	92	0.016	87	0.017	92
浙江	0.010	60	0.010	60	0.011	60	0.011	63
安徽	0.017	105	0.018	105	0.019	105	0.019	103
福建	0.010	60	0.011	65	0.013	70	0.012	68
江西	0.018	110	0.019	110	0.018	100	0.018	99
山东	0.024	149	0.027	159	0.031	169	0.027	151
河南	0.022	140	0.026	150	0.028	150	0.027	147
湖北	0.019	120	0.022	130	0.024	130	0.023	127
湖南	0.022	140	0.022	129	0.024	129	0.023	126
广东	0.002	15	0.003	20	0.004	23	0.005	29
广西	0.005	30	0.005	30	0.006	30	0.008	42
海南	0.009	58	0.010	58	0.011	58	0.010	57
重庆	0.005	30	0.005	30	0.006	30	0.007	36
四川	0.018	110	0.021	125	0.022	120	0.021	115
贵州	0.013	80	0.014	80	0.015	80	0.014	79
云南	0.010	60	0.010	60	0.011	60	0.011	63
陕西	0.033	210	0.036	210	0.039	210	0.036	196
甘肃	0.005	30	0.004	25	0.005	27	0.006	33
青海	0.005	32	0.001	5	0.001	5	0.003	15
宁夏	0.004	25	0.004	25	0.005	25	0.005	28
新疆	0.002	10	0.002	10	0.002	10	0.004	22
总计		6276		5864		5423		5520

注: □为招生计划增长型省份; □为招生计划减少型省份。

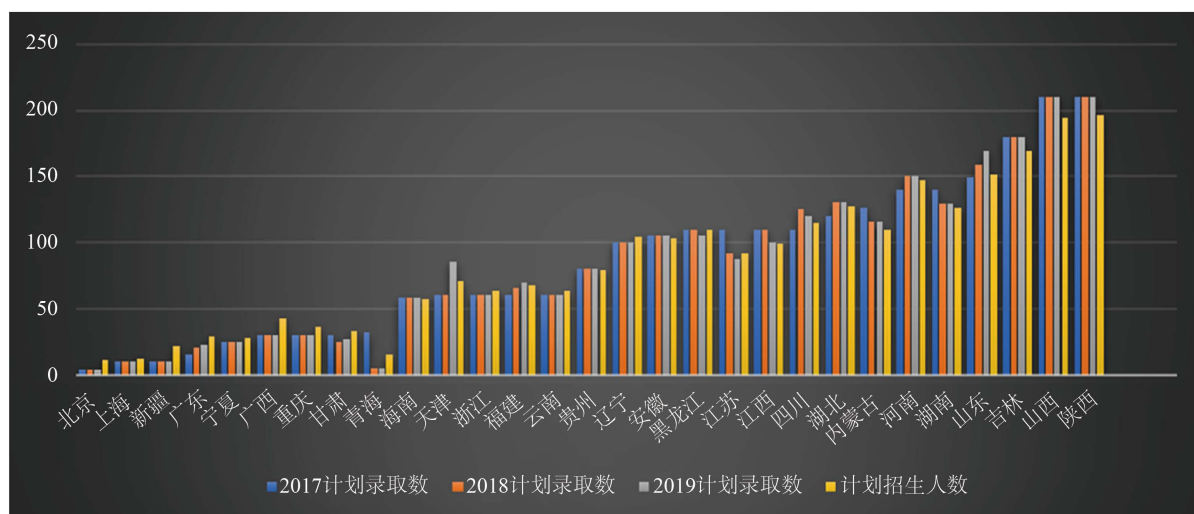


Figure 1. The chart of enrollment by province from 2017 to 2019

图 1. 2017~2019 年各省份招生情况条形图

结合表 4 和图 1 可以发现:

在利用熵值法计算的 2020 年招生计划中,北京、辽宁、上海、浙江、广东、广西、重庆、云南、甘肃、宁夏和新疆 11 个地区招生计划数比 2017~2019 年都高,说明熵值法决策中,这 11 个地区应扩大招生人数。以北京为代表,结合表 5 各指标排名情况可以发现,北京的历年招生计划比例最小,但招生质量和招生数量指标中,三个指标水平都很高,其中  $I_2$  表示在北京录取最低分与控制线差值,该值明显大于其他省份,说明在北京录取的考生质量更优秀,结合  $I_3$  和  $I_4$  招生数量,更加说明 2017~2019 年北京招生情况十分优秀,因此适当扩大该省份的招生分配名额是合理的。

山西、内蒙古、吉林、安徽、江西、湖南、海南、贵州和陕西 9 个地区招生计划数比 2017~2019 年都低,说明熵值法决策中,这 9 个地区应减少招生人数。以山西省为代表,其历年招生比例排名第二,仅次于河北,但招生质量和招生数量指标中只有  $I_4$  排在前十名,其余排名较为靠后,说明其招生情况相对其历年招生比例排名第二情况并不乐观,因此适当减少该省份招生计划名额是合理的,有助于平衡招生质量和数量与招生计划之间的关系,使各省份分配更合理。

天津、河北、黑龙江、江苏、福建、山东、河南、湖北、四川和青海 10 个地区招生计划数位于 2017~2019 年招生计划数之间,熵值法决策给出了更为合理的招生计划分配。这 10 个地区在 2017~2019 年的招生计划中是上下变动的,这可能是招生部门结合历年招生情况根据经验做出了一定的调整,具有较强主观性,无理论依据,而本文利用熵值法给出更为合理的分配方法,更加客观,具有理论依据支撑。以招生人数最多的河北省为例,2017~2019 年计划招生人数一直在降低,这符合招生质量和数量排名较低情况,但河北省作为主要生源来源地,显然 2019 年计划数过低,导致招生分配不合理,因此应在最高招生计划与最低招生计划数之间找到更为合理的招生计划值,熵值法利用历史数据给出了这一确定值更为合理。

Table 5. The ranking of each province in each index and the total weight

表 5. 各省份在各指标和总权重下的排名

省份	$I_1$	$I_2$	$I_3$	$I_4$	$I_5$	$W_i$
北京	30	4	27	1	8	10
天津	17	5	30	2	2	17
河北	1	15	22	25	27	24

## Continued

山西	2	11	17	15	4	3
内蒙古	9	25	12	30	1	30
辽宁	14	28	26	26	24	16
吉林	4	26	19	6	7	4
黑龙江	11	12	13	17	28	28
上海	28	27	11	3	3	1
江苏	15	19	7	14	16	6
浙江	19	7	20	7	26	2
安徽	13	8	5	5	22	23
福建	18	18	21	12	6	12
江西	12	14	29	8	23	9
山东	5	30	2	16	5	11
河南	6	3	28	10	21	21
湖北	8	6	4	20	12	13
湖南	7	22	9	9	9	29
广东	26	20	1	4	17	7
广西	22	9	25	13	25	19
海南	21	24	16	23	19	25
重庆	23	16	3	21	20	22
四川	10	21	24	11	10	15
贵州	16	2	10	29	15	5
云南	20	1	15	24	29	26
陕西	3	29	23	22	30	27
甘肃	24	17	6	19	11	20
青海	27	13	8	28	14	14
宁夏	25	10	14	18	18	8
新疆	29	23	18	27	13	18

本文利用熵值法对各省份招生计划进行合理分配,通过历史数据,给出各省份具体比例及招生计划值,并进一步将30个省份分为三大类进行具体分析,分别为以北京为代表的计划增长型11个地区,以山西省为代表的计划减少型9个地区和以河北为例的计划适当调整型10个地区,针对不同类型地区,招生部门在计划分配时应结合实际分情况对招生计划进行调整,使招生计划分配更合理,提高招生质量。

## 基金项目

本文得到了河北工程大学教育教学研究项目(JG2018023)的支持,在此表示感谢。

## 参考文献

- [1] 付辉. 基于灰色预测模型和层次分析法的高校招生数据分析与研究[D]: [硕士学位论文]. 重庆: 重庆交通大学,



---

2013.

- [2] 王江曼. 基于层次分析法的高校分省分专业招生计划编制研究[J]. 价值工程, 2017, 36(34): 189-192.
- [3] 郑超美. 浅论普通高校招生来源计划的科学编制[J]. 安徽工业大学学报(社会科学版), 2006(6): 131-133.
- [4] 杨卫平, 段丹青, 陈松岭. 基于模糊数学的高校招生来源计划编制模型研究[J]. 计算机应用研究, 2009, 26(2): 580-582.
- [5] 郑庆华, 罗京, 王衍波, 杨松, 宋红霞. 普通高校分省招生计划编制模型研究[J]. 计算机应用研究, 2012, 29(7): 2567-2570.
- [6] 陈琳. 基于高校招生需求的计划分配模型研究[J]. 西南师范大学学报(自然科学版), 2015, 40(7): 136-142.
- [7] 魏针, 李登峰. 基于破产理论的省属高校招生计划分配方法[J]. 福州大学学报(自然科学版), 2020, 48(1): 14-19.