

非参数统计在高校实力差异分析中的应用

张秋桥 厦门大学经济学院计划统计系

【摘要】本文首先考察了中国管理科学研究院给出的2010《中国大学评价》中大陆前100位不同类别高校的综合得分分布特征;然后,分别应用非参数统计分析中Kruskal-Wallis检验和Mann-Whitney检验对不同类别高校的相关评价得分进行了分析,以期发现不同类型高校在综合实力、人才培养、科学研究上的差异,并提出了提升不同类别高校实力的建议。

【关键词】高校评价 高校类型 非参数检验

本文利用非参数检验的方法,通过高校综合得分截面数据对我国部分高校综合实力与院校类别相关性问题展开讨论。非参数统计是统计学的一个重要分支,它与总体究竟是怎样的分布几乎没有关系,仅需要一些非常一般性的假设,进行统计推断时只利用样本观察值中一些非常直观的信息,所以非参数统计分析含有丰富的统计思想并在实践中有着广泛的应用。

一、高校综合得分指标、类型划分及非参数检验方法

1. 高校综合得分指标体系。对高校综合实力进行评价有诸多不同的指标。国际上广泛流行的是美国《基本科学指标》(ESI)数据库。国内有较大影响力的是中国管理科学研究院武书连连续14年发布的《中国大学评价》。其指标分布如下[2]:

表1 中国大学评价指标体系

一级指标	二级指标	三级指标
人才培养	本科生培养	本科毕业生就业率、新生录取分数线、本科教学评估结果、全校师生比、本科毕业生数、教师平均学术水平、双语教学示范课程、实验教学示范中心、特色专业、教学团队、规划教材、挑战杯本科生学术竞赛奖、本科数学建模竞赛奖、本科教学成果奖
	研究生培养	毕业生平均学术水平、博士毕业生数、硕士毕业生数、优秀学位论文、研究生教学成果奖
科学研究	自然科学研究	国内引文数据库论文及引用、国外引文数据库论文及引用、学术著作引用、艺术作品、专利授权、科学与技术奖、国家大学科技园
	社会科学研究	国内引文数据库论文及引用、国外引文数据库论文及引用、学术著作引用、艺术作品、专利授权、人文社会科学奖

计算大学总得分:

设被评价的大学数为 n , A_k 为 k 大学总得分, B_{1k} 为 k 大学人才培养得分, B_{2k} 为 k 大学科学研究得分, a_1 为人才培养权重系数, a_2 为科学研究权重系数。则有:

2. 我国高校的分类划分与修正。

2.1 传统分类。按照学科设置和办学特色,我国的高校大致可分为以下几类:

以北京大学、复旦大学、南京大学、厦门大学等为代表的综合性大学。这些大学历史悠久、理工文医学科门类齐全、文理渗透,文化底蕴较为深厚。

以清华大学、上海交通大学、西安交通大学、同济大学等为代表的以工科为主的综合性大学。这类大学多有着辉煌的历史,工科专业实力很强,为国家培养了大量高水平的优秀工程技术人员。

以各类工业、农林牧水产业、地矿、财经、政法、语言等行业培养专业人才为目的,由专业学院发展演变而成的专业学院式的准综合大学。这类院校多是在1952年院系调整后的特定历史条件下建立的,为数众多,教学水平参差不齐。

还有一类以北京师范大学、华东师范大学、东北师范大学、华中师范大学、华南师范大学和陕西师范大学为代表的师范类院校。这类院

校也属综合类大学。

2.2 必要的修正。笔者在研究时发现以下几个问题:

第一,第三类即专业学院式大学目前综合实力较为薄弱,进入《中国大学评价》前100的比例极少,故无法进行进一步的分析;

第二,一些大学虽然文理兼有,但其学科建设尚未达到综合性大学的标准,如重庆大学、湖南大学等;

第三,若干师范类、农林类大学学科设置及教学水平已超出其特定专业培养目的,可属于综合类、文理类大学。

基于以上思考,通过对比《中国大学评价》中的学校参考类型,将进入《中国大学评价》前100为的学校划分为:综合类、理科类、工科类和文理类。具体见下表:

表2 2010《中国大学评价》前10011所大学分类

综合类	理科类	工科类	文理类
清华大学	浙江大学	哈尔滨工业大学	吉林大学
北京大学	上海交通大学	东南大学	西安交通大学
复旦大学	华中科技大学	天津大学	南开大学
南京大学	中国科学技术大学	大连理工大学	北京师范大学
武汉大学	中南大学	同济大学	重庆大学
中山大学	兰州大学	华南理工大学	华东师范大学
四川大学	中国农业大学	北京航空航天大学	湖南大学
山东大学	西北农林科技大学	西北工业大学	南京农业大学
中国人民大学	中国海洋大学	北京理工大学	西南大学
厦门大学	中国地质大学	华东理工大学	东北师范大学
郑州大学	福州大学	南京航空航天大学	南京师范大学
上海大学	中国石油大学	电子科技大学	华中师范大学
苏州大学	贵州大学	武汉理工大学	东北大学
暨南大学		西安电子科技大学	西南交通大学
西北大学		南京理工大学	华南师范大学
扬州大学		北京科技大学	湖南师范大学
湘潭大学		北京交通大学	华中农业大学
河北大学		中国矿业大学	南昌大学
黑龙江大学		东华大学	陕西师范大学
		北京化工大学	河海大学
		江南大学	山西大学
		北京工业大学	河南大学
		合肥工业大学	云南大学
		燕山大学	华南农业大学
		浙江工业大学	首都师范大学
		江苏大学	山东农业大学
		华北电力大学	上海师范大学
		南京工业大学	安徽师范大学
		哈尔滨工程大学	广西大学

实证分析

综合类	理科类	工科类	文理类
		北京邮电大学	山东师范大学
		长安大学	

3. 非参数检验方法的适用。非参数检验中的 Kruskal - Wallis 检验和 Mann - Whitney 检验可分别对多样本和两样本差异进行检验。Kruskal - Wallis 检验和 Mann - Whitney 检验的一般提法可参见王静龙、梁小筠编著《非参数统计分析》。

二、实证研究

1. 不同类别高校综合得分的分布特征。本文数据来源于中国管理科学研究院武书连给出的 2010《中国大学评价》中的高校综合得分数据。利用上述指标评价体系对教育部、国家统计局发布的中国教育统计年鉴、中国科技统计年鉴中的数据计算而来。我们选取评价前 100 的大学综合得分来考察高校不同类别的差异。

表 3 不同类别高校综合得分的描述性统计分析

高校类别	样本数	均值	中位数	标准差	峰度	偏度	正态性检验 P 值
综合类	19	71.0	54.8	54.6	0.36	1.05	0.015
理科类	13	63.1	44.8	55.7	0.97	1.34	0.014
工科类	31	36.54	29.94	18.77	1.42	1.31	<0.005
文理类	30	33.77	25.86	20.20	1.74	1.56	<0.005

依据每个类别的得分,分别计算均值、中位数、标准差、峰度和偏度,并做 Anderson - Darling 正态性检验,结果列于表 3,由表 3 可知,综合类、理科类、工科类、文理类高校得分均值依次递减,且正态性检验 P 值都小于 0.05,所以其分布不服从正态分布。

2. 不同类别高校综合得分的非参数检验。由于高校得分分布不服从正态分布,所以考虑用非参数检验的方法来考察其差异。对于多样本的比较问题,可以采用 Kruskal - Wallis 检验。检验结果如下:

表 4 在不同类别高校综合得分上的 Kruskal - Wallis 检验

类别	N	中位数	平均秩	Z
综合类	19	54.84	60.3	2.42
理科类	13	44.80	52.2	0.74
工科类	31	29.94	44.9	-0.53
文理类	30	25.86	38.5	-2.10
整体	93		47.0	

H = 8.30DF = 3P = 0.040
H = 8.30DF = 3P = 0.040(已对结调整)

由表 4 可知,检验统计量 $H = 8.30$, 检验的 P 值为 $P(\chi^2(3) \geq 8.30) = 0.04 < 0.05$, P 值够小,所以认为不同类别高校的综合得分在显著性水平为 0.05 上存在显著性差异。

直观上看,综合类、理科类、工科类、文理类高校得分均值依次递减,若两两作 Mann - Whitney 检验,得出以下结果:

表 5 两两类别高校得分的 Mann - Whitney 检验

	(综合类,理科类)	(理科类,工科类)	(工科类,文理类)
ETA1 - ETA2 的点估计	5.90	7.66	3.01
W 值	332.0	327.0	1033.0
结论	在 0.2449% 上, ETA1 = ETA2 与 ETA1 > ETA2 的检验结果显著	在 0.1909% 上, ETA1 = ETA2 与 ETA1 > ETA2 的检验结果显著	在 0.1512% 上, ETA1 = ETA2 与 ETA1 > ETA2 的检验结果显著
	在 0.2449% 显著性水平上, 检验结果显著(已对结调整)	在 0.1909% 显著性水平上, 检验结果显著(已对结调整)	

由表 5 可知,综合类高校得分 > 理科类高校得分 > 工科类高校得

分 > 文理类高校得分均通过了显著性检验,因此上述直观上得出的结论是正确的。

3. 不同高校类别一级指标得分的非参数检验。高校综合得分是由人才培养和科学研究两个一级指标得分加权计算得来。下面对两个一级指标得分进行 Kruskal - Wallis 检验。检验结果如下:

表 6 在不同类别高校人才培养一级指标得分上的 Kruskal - Wallis 检验

类别	N	中位数	平均秩	Z
综合类	19	28.80	60.4	2.42
理科类	13	20.76	50.7	0.53
工科类	31	17.30	44.4	-0.67
文理类	30	15.52	39.7	-1.81
整体	93	47.0		

H = 7.42DF = 3P = 0.060

表 7 在不同类别高校科学研究一级指标得分上的 Kruskal - Wallis 检验

类别	N	中位数	平均秩	Z
综合类	19	26.04	59.5	2.27
理科类	13	24.04	55.1	1.16
工科类	31	12.93	44.4	-0.66
文理类	30	10.81	38.3	-2.16
整体	93	47.0		

H = 8.70DF = 3P = 0.034
H = 8.70DF = 3P = 0.034(已对结调整)

由表 6 可得,检验统计量 $H = 7.42$, 检验的 P 值为 $P(\chi^2(3) \geq 7.42) = 0.06 > 0.05$, P 值不够小,所以认为不同类别高校人才培养一级指标得分在显著性水平为 0.05 上不存在显著性差异。

由表 7 可知,检验统计量 $H = 8.70$, 检验的 P 值为 $P(\chi^2(3) \geq 8.70) = 0.034 < 0.05$, P 值够小,所以认为不同类别高校在科学研究一级指标得分在显著性水平为 0.05 上存在显著性差异。

从综合得分、人才培养一级指标得分、科学研究一级指标得分的 Kruskal - Wallis 检验结果来看,其 P 值分别为 0.04、0.06、0.034。因此可以认为不同类型高校的得分差异主要是体现在科学研究上。而且从科学研究指标得分的中位数来看,综合类、理科类高校要明显高于工科类、文理类高校。

三、结论与建议

本文通过非参数检验的方法,对我国大陆不同类型高校的相关评价得分进行了分析,发现不同类型的高校确实存在着相关评价得分的差异。第一,将高校分为综合类、理科类、工科类、文理类,在综合得分上,四类高校存在明显差异。且得分由高到低分别为综合类、理科类、工科类、文理类。第二,从人才培养和科学研究两个一级指标的得分来看,在人才培养上,四类高校并不存在显著性差异,而在科学研究上,则差异显著。且综合类、理科类高校要明显高于工科类、文理类高校。

不同类型的高校在学科设置、教学水平上都有各自特色。从本文的结论来看,第一,工科类、文理类高校应调整办学思路与模式,完善学科建设和专业设置,适应社会需要,在保存和发展自身特色的基础上,逐渐向综合性大学发展。第二,不同类型高校在科学研究上应加强相互沟通与合作,尤其是工科类和文理类高校,应重视科研能力的提升,以此强大办学力量。

参考文献:

- [1]王静龙,梁小筠.非参数统计分析.北京:高等教育出版社[M].2006:94-96.
[2]武书连,吕嘉,郭石林.2010 中国大学评价[J].科学管理研究.2010(4):5-9.