

# 中国优秀博士学位论文计量分析

邱均平,赵月华

(武汉大学,湖北 武汉 430072)

**摘要:**运用计量的方法和模型,从学位授予单位、学科和论文指导教师三个方面对中国优秀博士学位论文评选结果进行统计和分析,并得出以下结论:(1)优秀博士的培养已经形成了由“985工程”高校和中国科学院共同组成的核心群体;(2)理工类共占优秀博士学位论文总数的67.79%,显示出学科发展不均衡、结构不合理等问题;(3)指导教师与其培养出的优秀博士学位论文篇数服从洛特卡分布: $f(y_x) = 0.759/x^{3.302}$ ,并且各学科门类历年获奖论文导师与学位办评议组成员存在高度相关性,Pearson(皮尔生)乘积矩相关系数高达0.974,在一定程度上揭示出现行评选机制存在的问题。

**关键词:**博士学位论文;优博;计量分析;洛特卡分布;相关性分析

**中图分类号:**C961

**文献标志码:**A

**文章编号:**1008-5831(2014)02-0092-07

## 一、引言

作为“高层次创造性人才工程”重要组成部分的全国优秀博士学位论文(以下简称优博论文)的评选工作,已纳入教育部《面向21世纪教育振兴行动计划》的第一项启动项目,全国优秀博士学位论文评选是教育部学位管理与研究生教育司组织开展的一项工作,旨在加强高层次创造性人才的培养,鼓励创新精神,提高研究生教育特别是博士生教育的质量。中国优秀博士学位论文评选是对博士培养质量进行监督和激励的一项重要举措,对培养和激励创新精神,促进博士生培养质量提高具有积极的作用<sup>[1]</sup>。

根据《全国优秀博士学位论文评选办法》,全国优秀博士学位论文入选名单经过推荐、初选和复评后产生。参评论文由学位授予单位向其所在省(自治区、直辖市)学位委员会推荐;省级学位委员会对推荐学位论文进行初选。教育部学位管理与研究生教育司负责组织对初选出的论文进行复评,复评工作包括同行专家通讯评议和专家审定会审定。教育部学位与研究生教育发展研究中心受教育部学位管理与研究生教育司委托,承担同行专家通讯评议工作。

## 二、历年中国优秀博士学位论文评选结果数据统计分析

自1999年首次进行中国优秀博士学位论文评选以来,至2011年共进行了13次,共评选出中国156个高校/科研单位的优秀博士学位论文1279篇,提名论文1902篇,如表1所示<sup>[2]</sup>。

根据《全国优秀博士学位论文评选办法》中的说明,评选工作每年进行一次,每次评选出的全国优秀博士学位论文不超过100篇。评选工作进行13年来,评选结果平均为每年98.4篇,说明评选工作符合“科学公正、注重创新、严格筛选、宁缺毋滥”的原则。

2010年全国毕业博士生4.90万人,评选出97篇全国优秀博士学位论文,仅占总数的0.2%。在评选第一年1999年,参评论文基数仅为8957篇,评选出的100篇优秀论文占总数的1.1%。可见,虽然随着国家博士生的逐渐扩招,每年的博士毕业论文数量急剧增加,2011年全国博士生招生达到6.38万人,然而,论文评选依然不超过100篇,使得获评难度急剧增大,获评几率大幅降低。

收稿日期:2013-09-29

**作者简介:**邱均平(1947-),男,湖南涟源人,武汉大学信息管理学院教授,博士研究生导师,武汉大学中国科学评价研究中心主任,《评价与管理》杂志主编,主要从事信息计量与科学评价研究;赵月华,武汉大学信息管理学院管理科学与工程专业硕士研究生,主要从事信息管理及科学评价研究。

表1 历年全国优秀博士学位论文评选结果

年份	获奖论文	提名论文	年份	获奖论文	提名论文
1999	100	无	2006	99	137
2000	100	无	2007	98	158
2001	100	无	2008	100	177
2002	97	无	2009	98	363
2003	97	179	2010	100	334
2004	97	139	2011	97	256
2005	96	159			

全国优秀博士学位论文评选的初衷在于加强高层次创造性人才的培养工作,鼓励创新精神,然而可以预见,未来每年的参评论文基数将大幅上升,获评论文将低于总数的0.2%,显然这将大大降低参评博士生的热情,无法起到鼓励人才的作用。因此,建议能够随着博士生毕业人数的增长,将获奖论文的数量控制在1%左右,保持与评选初年持平的水平。

### 三、历年获奖论文学位授予单位分析

#### (一) 学位授予单位分析

中国优秀博士学位论文的参评单位包括全国所有具有博士学位授权资格的单位,即高校、中国科学院、中国社会科学院以及其他部委所属科研院所等。截止2010年底,全国共有培养研究生单位797个,其中高等学校481个,科研机构316个。至2011年,1279篇获奖论文共来自156个高校/科研单位,对其中“985工程”、“211工程”高校的获奖论文数进行比较分析,可以一定程度上反映中国高等教育政策对高校研究生教育和科研工作的影响效果。获奖总数排在前十位的单位如表2所示。

表2 获奖总数排名前十位的学位授予单位

学位授予单位	篇数	占总数的百分比
中国科学院	184	14.39%
清华大学	103	8.05%
北京大学	85	6.65%
复旦大学	54	4.22%
浙江大学	44	3.44%
中国科学技术大学	39	3.05%
南京大学	37	2.89%
上海交通大学	31	2.42%
中国人民大学	27	2.11%
中山大学	25	1.95%

中国科学院共有44个下属科研院所获评优博论文,总数达184篇,占获奖论文总数的14.39%,并且相比排在第二和第三位的清华大学和北京大学,其获奖总数都在2倍左右。分析其原因,一方面因为中国科学院博士生培养规模庞大,2011年共有4832人获得博士学位,相比其他高校和科研单位,评选基数较大;另一方面,中国科学院拥有320余位两院院士、4200余名博士生导师,其以研究所为基础的高水平导师队伍和一流科研实践条件,为培养优秀博士生提供了肥沃的土壤。

在排名前十位的单位中,除中国科学院外,全部

是国家“985工程”一期大学。这一方面说明“985工程”一期建设高校都是国内高水平研究型大学;另一方面,也体现了“985工程”大学依托其优势学科创新平台项目,为国家高精尖人才培养做出了卓越的贡献。

在156个学位授予单位中,包含了39个“985工程”高校(共39所),和89个“211工程”高校(共113所),也就是说,包括了全部“985工程”高校以及78.76%的“211工程”高校,39所“985工程”高校共产出优博论文768篇,加上中国科学院的184篇,这40所高校/科研单位共产出952篇优博论文,相当于25.64%的学位授予单位产出了74.43%,大致符合“二八率”的分布,说明中国优秀博士生的培养已经形成了由“985工程”高校和中国科学院共同组成的核心群体。

#### (二) 学位授予单位的洛特卡分布统计

中国优博论文的评选对于衡量一个单位(高校或科研机构)的研究生教育和科研水平具有重要作用,“优博生产率”可以作为衡量研究生教育水平的重要指标。因此通过对获优论文的单位分布进行模型拟合,可以反映出中国博士教育和培养的发展现状。

对13年来的数据进行统计,在156个高校/科研单位中,平均获优数是8.2篇,样本标准差为19.28篇,离散系数为2.35,极差为183篇,这表明中国博士培养单位的科研水平相差距离很大。

通过对历年获奖论文进行统计,设 $x$ 为高校/科研单位的优博论文篇数, $y_x$ 为产出 $x$ 篇优博论文的高校/科研单位数量, $f(y_x) = y_x / \sum y_x$ 。为了便于观察结果,将 $x$ 与 $y_x$ 占高校/科研单位总数的比例 $f(y_x)$ 分别取对

数,得到  $lgx$ 、 $lgf(y_x)$ , 然后根据所得数据绘制出获优论文单位的洛特卡分布图,如图 1 所示。

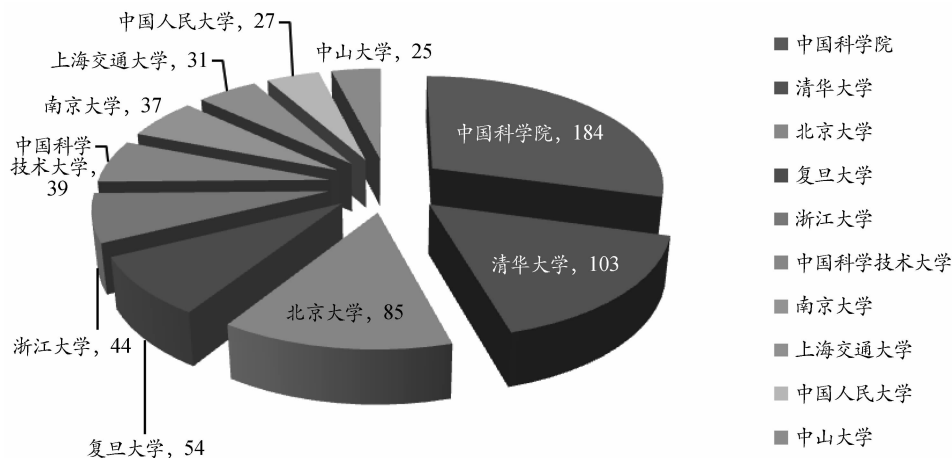


图1 历年获优论文单位的洛特卡分布图

利用 Excel 软件对样本数据进行最小二乘法参数估计,构建出优博论文篇数  $x$  与产出  $x$  篇优博论文的高校/科研单位数量  $y_x$  占高校/科研单位总数的比例  $f(y_x)$  的 Lotka 分布数学模型:

$$\ln f(y_x) = -0.822 \ln x - 1.991$$

此模型的拟合优度较高 ( $R^2$  值达到 0.710),表明“优博生产率”频率分布符合广义上的洛特卡分布,得到中国优秀博士学位论文与高校/科研单位频率的 Lotka 分布为:

$$f(y_x) = 0.137/x^{0.822}$$

#### 四、历年获奖论文学科分布分析

##### (一) 历年获奖论文学科门类分布分析

1 279 篇优博论文分布在 13 个学科门类(图 2),分布在工学、理学和医学三个学科门类的论文数排在前三名,分别占获奖论文总数的 38.70%、29.09% 和 9.38%。理工类共占总数的 67.79%,超过 2/3,明显高于其他学科门类,反映了中国经济社会发展在向工业化、信息化转型时期的知识需求和产业需要。但是,这种过于集中的获奖分布,也显示出中国学科发展不均衡、结构不合理等问题,不利于国家整体科研环境的良性发展,同时,也在一定程度上阻碍了人文社科类学科的发展。

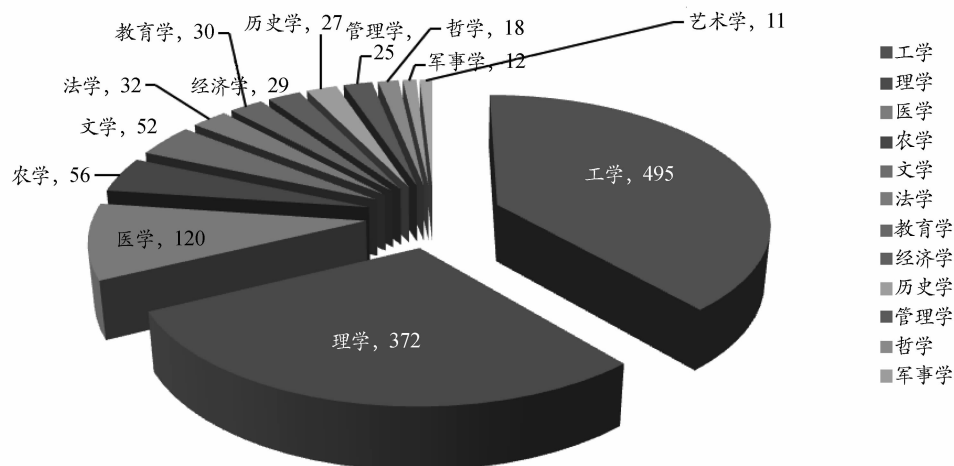


图2 获奖论文学科门类分布

##### (二) 历年获奖论文一级学科分布分析

在 115 个博士招生一级学科中(包括 5 个博士专业学位),共有 85 个一级学科获得优博论文,根据国务院学位委员会、教育部印发的《学位授予和人才培养学科目录设置与管理办法》中公布的《学位授予和人才培养学科目录》,对有变动的学科进行调整对应后,统计出共有 21 个一级学科从未获得过优博论文。

获奖数排在前十名的一级学科及其优博论文总篇数如图 3 所示。化学、物理学和生物学分别以 86 篇、

82 篇和 77 篇高居前三位,并且都属于理学,充分显示出理学基础学科是中国的优势学科,凸显出学科发展的不平衡。

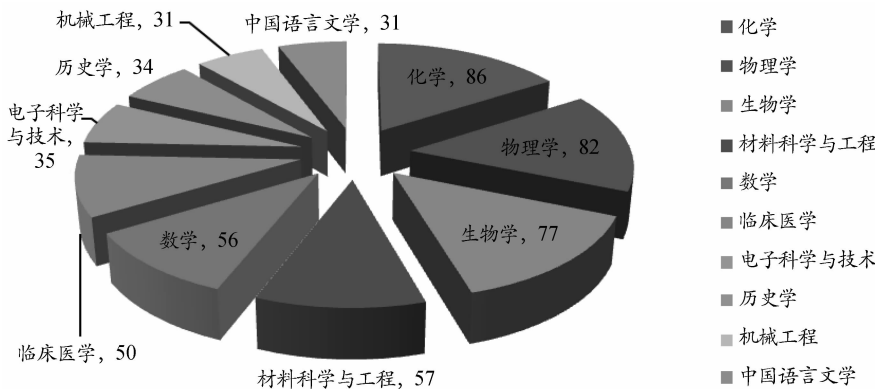


图3 获奖论文一级学科分布

五、历年获奖论文导师分析

自 1999 年首次评选以来,共有 1 067 位博士生导师指导的论文获得优博论文。对历年评选结果进行统计,发现培养出 10 篇优博论文的指导教师有 1 人,5 篇有 1 人,4 篇有 6 人,3 篇有 28 人,2 篇有 125 人,1 篇有 906 人。篇数大于等于 4 篇的导师姓名、一级学科和学位授予单位如表 3 所示。

表3 获奖总数大于等于 4 篇的指导教师

指导教师姓名	篇数	一级学科	学位授予单位
曹雪涛	10	基础医学	第二军医大学/浙江大学
李德仁	5	测绘科学与技术	武汉大学
樊代明	4	临床医学	第四军医大学
郭光灿	4	物理学	中国科学技术大学
贺林	4	生物学	上海交通大学
裴钢	4	生物学	中国科学院
汤钊猷	4	临床医学	复旦大学
张大鹏	4	园艺学	中国农业大学

表4 历年获奖论文导师的分布

x	$y_x$	$f(y_x)$
10	1	0.09%
5	1	0.09%
4	6	0.56%
3	28	2.62%
2	125	11.72%
1	906	84.91%

(一) 历年获奖论文导师的洛特卡分布统计

1926 年,洛特卡以其独特的视角,阐述了期刊论文作者与论文数量之间的定量关系——洛特卡平方反比率,并引入了“科学生产率”的概念,即“个体科研人员在一定时期内所撰写的论文数量”<sup>[3]</sup>,通过这一概念的引入来测量科研人员撰写科学文献的能力。对于博士生导师来说,培养出优博论文的数量,即“优博生产率”,可以用来测量博士生导师培养优秀博士生的能力。

通过对历年获奖论文导师进行统计,设  $x$  为博士生导师培养出的优博论文篇数,  $y_x$  为培养出  $x$  篇优博论文的博士生导师数量,  $f(y_x) = y_x / \sum y_x$ 。依次计算出的样本数据表 4 所示。

为了便于观察结果,将优博论文篇数  $x$ , 与博导数  $y_x$  占博导总数的比例  $f(y_x)$  分别取对数,得到  $lgx, lg f(y_x)$ , 然后根据所得数据绘制出获奖论文导师的洛特卡分布图,如图 4 所示。利用 Excel 软件对样本数据进行最小二乘法参数估计,构建出优博论文篇数  $x$  与博导数  $y_x$  占博导总数的比例  $f(y_x)$  的 Lotka 分布数学模型:

$$\ln f(y_x) = -3.302 \ln x - 0.276$$

此模型的拟合优度很高 ( $R^2$  值达到 0.916),表明“优博生产率”频率分布符合广义上的洛特卡分布,得

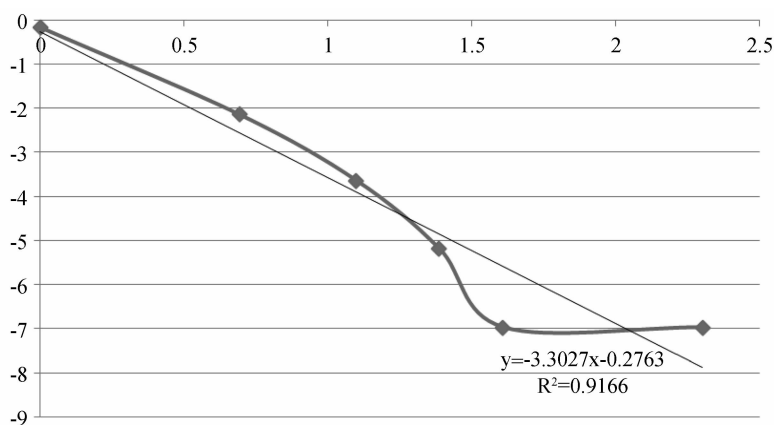


图4 历年获奖论文导师的洛特卡分布

到全国优秀博士学位论文与博导频率的 Lotka 分布为:

$$f(y_x) = 0.759/x^{3.302}$$

从图4可以看出,有两个点距离构建出的洛特卡分布的拟合直线模型较远,这两个点分别为x值为5和10的点,绝对偏差分别达到1.382和0.907。从原始数据中发现,产出5篇和10篇优博论文的分别只有1位博导,他们属于洛特卡所定义的“杰出者”。

“优博生产率”最高的是13年来产出10篇优博论文的曹雪涛院士。他是免疫学专家,现任第二军医大学副校长、第二军医大学免疫学研究所所长、博士生导师,2005年当选为中国工程院院士<sup>[4]</sup>。由他领军的免疫学教研室,在肿瘤的免疫和基因治疗领域取得了一系列具有应用价值的成果,其中最为突出的是发现了一种具有独特调控功能的新型树突状细胞亚群,其研究成果2004年在国际顶级学术杂志《自然·免疫学》上以封面论文的形式发表。他所指导的优博论文中,绝大多数都是对树突状细胞的研究。说明博士生导师的重大研究成果将在很大程度上对博士生的学习起到铺垫和引领的作用。

#### (二) 历年获奖论文导师与学位办评议组成员相关性分析

根据《全国优秀博士学位论文评选办法》第十一条规定,教育部研究生工作办公室负责组织对初选出的论文进行复评,复评工作包括同行专家通讯评议和专家审定会审定。将获得优博论文的指导教师与1999年至2011年间国务院学位委员会学科评议组成员名单进行对比(该时间段涉及第四届、第五届和第六届学科评议组),发现24.27%的指导教师担任国务院学位委员会学科评议组成员。将每年获奖的指导教师与当年的学科评议组成员名单对比,统计出优博论文指导老师中担任评议组成员的比重,并将各年比重的变化绘制成曲线图,如图5所示。

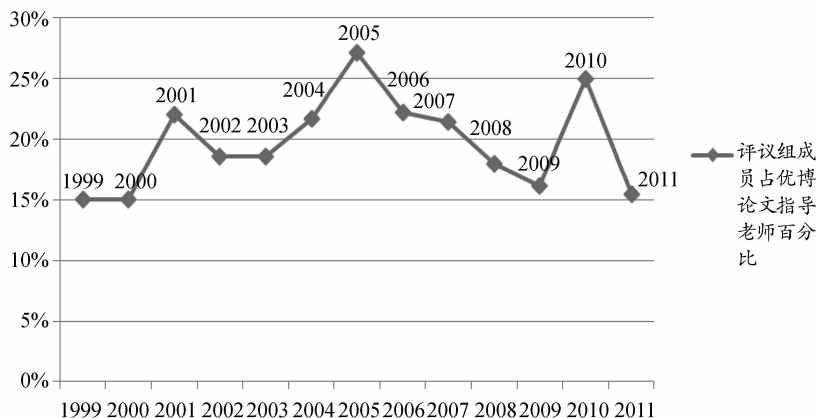


图5 优博论文指导老师中担任评议组成员比重变化图

在比重最高的2005年,27.08%获得优博论文的指导教师是当年的评议组成员,即使是最低的1999年和2000年,优博论文指导老师中担任评议组成员的比重也达到15%。

特别值得注意的是,在36名高产指导教师中(指导出大于等于3篇优博论文),产出10篇的曹雪涛和产出5篇的李德仁都是评议组成员,另外还有18名都在当年担任评议组成员,高产指导教师中评议组成员所

占比重高达 55.56%。

### (三) 各学科门类历年获奖论文导师与学位办评议组成员相关性分析

根据国务院学位委员会、教育部印发的《学位授予和人才培养学科目录设置与管理办法》中公布的《学位授予和人才培养学科目录》中的 13 个学科门类,对 1 297 篇获奖论文进行统计归类,并将学位委员会学科评议组成员指导的论文进行统计。如表 5 所示。

表 5 各学科门类历年获奖论文导师与学位办评议组成员相关性

学科门类	总篇数	评议组成员 指导的论文数	百分比
管理学	25	14	56.00%
教育学	30	16	53.33%
哲学	18	9	50.00%
医学	120	48	40.00%
法学	32	11	34.38%
工学	495	142	28.69%
艺术学	11	3	27.27%
文学	52	14	26.92%
农学	56	14	25.00%
理学	372	76	20.43%
军事学	12	2	16.67%
历史学	27	4	14.81%
经济学	29	4	13.79%

将结果输入 Excel 软件,选择 Pearson(皮尔生)函数对两组数据进行相关性分析。Pearson(皮尔生)乘积矩相关系数  $r$  的公式为:

$$r = \frac{\sum (x - \bar{x})(y - \bar{y})}{\sqrt{\sum (x - \bar{x})^2 \sum (y - \bar{y})^2}}$$

其中  $\bar{x}$  和  $\bar{y}$  分别是两组数据的平均值,得到的相关系数  $r$  是一个范围在 -1.0 到 1.0 之间(包括 -1.0 和 1.0 在内)的无量纲指数,反映了两个数据集之间的线性相关程度。对表 5 中的数据进行计算,得到相关系数  $r$  为 0.974,按照 Pearson(皮尔生)函数的说明,  $|r|$  大于等于 0.8 时,可以认为两变量间高度相关。因此,可以认为各学科门类历年获奖论文导师与学位办评议组成员存在高度相关性。

统计结果显示,13 个学科门类的获奖论文指导

教师与评议组成员相关性平均为 31.33%,相关性超过 50% 的学科门类为管理学、教育学和哲学。管理学和哲学的总篇数分别排在各学科门类论文总数的第 10 位和第 12 位,但是其相关性却排在第 1 位和第 3 位。说明这两个学科门类的研究成果并不显著,并且仅有的少量成果较高程度地集中于评议组成员指导的博士。

表 6 各一级学科历年获奖论文导师与学位办评议组成员相关性

一级学科名称	总篇数	评议组成员 指导的论文数	百分比
公共管理	2	2	100.00%
军制学	1	1	100.00%
社会学	1	1	100.00%
图书馆、情报与档案管理	2	2	100.00%
交通运输工程	9	7	77.78%
林业工程	4	3	75.00%
新闻传播学	4	3	75.00%
基础医学	26	18	69.23%
农业工程	6	4	66.67%
农林经济管理	5	3	60.00%
轻工技术与工程	5	3	60.00%
教育学	12	7	58.33%
心理学	14	8	57.14%
管理科学与工程	9	5	55.56%
兽医学	6	3	50.00%
畜牧学	8	4	50.00%
战役学	2	1	50.00%
哲学	18	9	50.00%
中西医结合	4	2	50.00%
中药学	2	1	50.00%

### (四) 各一级学科历年获奖论文导师与学位办评议组成员相关性分析

将历年获奖论文的数据,按照一级学科进行统计,统计出各一级学科获奖论文的总篇数,以及其中指导老师是学位评议组成员的文章数量。同样将数据导入 Excel 软件,选择 Pearson(皮尔生)函数对两组数据进行相关性分析,得到相关系数  $r$  为 0.836,大于 0.8,说明各一级学科历年获奖论文导师与学位办评议组成员同样呈现出高度的相关性。

将其中指导教师是学位评议组成员的文章数量与总篇数进行比重分析,按照所占比重降序排列,比重大于等于 50% 的学科如表 6 所示。

在 85 个获得优博论文的一级学科中,历年获奖论文导师与学位办评议组成员相关性平均为 32.74%,也就是说,对于所有一级学科而言,将近 1/3 的获奖论文来自评议组成员指导的博士。并且,超过 50% 的一级学科有 20 个,有 4 个一级学科相关性达到 100%,分别为公共管理、军制学、社会学和图书馆、情报与档案管理。相比按照学科门类统计得到的相关性,以一级学科进行分类统计得到的相关性更加明显。显示出一些学科获奖论文成果高度集中在评议组成员指导的博士中。

通过以上分析可以看出,无论是对于学科门类或者一级学科,历年获奖论文导师与学位办评议组成员都呈现出高度的相关性。这在一定程度上说明了中国博士培养的学科发展的不均衡性。尤其是对于学位办评议组成员占获奖论文导师比重超过50%的一级学科,杰出的研究人员呈现出高度的集中性,形成了该领域的核心团体。然而,指导教师与评议组成员的高度相关性也反映出优势资源的高度集中,甚至在一定程度上抑制了百家争鸣的良好学术发展氛围。另一方面,中国优秀博士学位论文的评选结果已经普遍作为高校对博士生培养单位以及指导教师绩效评价的重要指标,那么如果能够建立更加有效的评选机制,避免指导教师与评议组成员高度的相关性,将能够大大增强评选的信度。

## 六、结论与思考

自1999年首次进行中国优秀博士学位论文评选以来,至2011年共进行了13次,评选出1279篇优博论文。运用计量的方法和模型,从学位授予单位、学科和论文指导教师三个方面进行统计和分析。

首先,从学位授予单位看,39所“985工程”高校和中国科学院共产出了74.43%的优博论文,大致符合“二八率”分布,说明中国优秀博士生培养已经形成由“985工程”高校和中国科学院共同组成的核心群体。

其次,就学科分布而言,理工类优秀博士学位论文共占总数的67.79%,反映了中国经济社会发展在向工业化、信息化转型时期的知识需求和产业需要,同时也显示出中国学科发展不均衡、结构不合理等问题。

最后,通过对指导教师进行统计分析, $x$ 为博士生导师培养出的优秀博士学位论文篇数, $y_x$ 为培养出 $x$ 篇优秀博士学位论文的博士生导师数量之间,服从洛特卡分布: $f(y_x) = 0.759/x^{3.302}$ 。进一步对各学科门类和一级学科历年获奖论文导师与学位办评议组成员相关性进行分析发现,相关系数分别高达0.974和0.836,高度的相关性说明杰出的研究人员呈现出高度的集中性,形成了该领域的核心团体。另一方面,也反映出现行评选机制使得优势资源过度集中,甚至在一定程度上抑制了百家争鸣的良好学术发展氛围。

中国优秀博士学位论文评选的初衷在于建立有效的监督和激励机制,提高博士生教育的整体质量,鼓励创新,促进高层次创新人才脱颖而出。然而通过以上统计和分析结果,反映出中国学科发展不均衡、资源配置不合理,以及优秀博士学位论文评选机制存在的问题,值得广泛的关注和思考。

## 参考文献:

- [1] 教育部学位与研究生教育发展中心. 全国优秀博士学位论文评选项目简介[EB/OL]. [2012-05-22]. <http://www.cdg-de.edu.cn/xwyyjsjyxx/zlpj/yblwpm/257695.shtml>.
- [2] 教育部学位与研究生教育发展中心. 历年全国优秀博士学位论文评选结果[EB/OL]. [2012-05-22]. <http://www.cdg-de.edu.cn/xwyyjsjyxx/zlpj/yblwpm/>.
- [3] 邱均平. 信息计量学[M]. 武汉:武汉大学出版社,2007:161.
- [4] 中国工程院. 工程科技界的杰出人才——曹雪涛[EB/OL]. [2012-05-22]. <http://www.cae.cn/cae/jsp/jump.jsp?oid=20111231115415843674246>.

## Quantitative Analysis of National Excellent Doctoral Dissertations

QIU Junping, ZHAO Yuehua

(Wuhan University, Wuhan 430072, P. R. China)

**Abstract:** Using quantitative method and model, from three aspects such as the degree conferring unit, discipline and the Dissertation instructor, this paper conducts the statistics and analysis of the selection results of the National Excellent Doctoral Dissertations, and draws the following conclusions: (1) Excellent Doctors have formed a core group of “985 Project” colleges and universities and China Academy of Sciences; (2) Science and engineering accounted for 67.79% of the total outstanding doctoral dissertations, showing the problem of uneven development of the discipline and the unreasonable structure; (3) instructors and the number of Excellent Doctoral Dissertations obey Lotka distribution:  $f(y_x) = 0.759/x^{3.302}$ , and the instructors of awarded Excellent Doctoral Dissertations highly correlates to Appraisal Group, Pearson correlation coefficient is 0.974, appearing the problems of selection mechanism to some extent.

**Key words:** excellent doctoral dissertations; quantitative analysis; Lotka distribution; correlation analysis