

卡方检验在学生成绩差异性分析中的应用

徐 向 阳

(南京工业大学 经济管理学院, 江苏 南京 210009)

摘 要: 用卡方检验方法对作业成绩、考试成绩和总成绩进行差异性分析, 有助于找出学生学风中存在的问题, 为提高教学水平提供有益的帮助。

关键词: 卡方检验; 学生成绩; 差异分析

中图分类号: G642.0 **文献标识码:** A

0 引 言

在高等院校教学过程中, 常常遇到这样的情况: 同一个教师在同一个课堂, 用同一本教材、同一张试卷, 所教的不同班级考试成绩却大相径庭。这种现象发人深思: 为什么考试成绩反差如此大? 为什么平时作业显示不出班级之间的学习差异, 而考试结果却大大出乎意料? 是考题出偏了、出难了? 还是学生学风有问题, 作业存在虚假性? 如何来定量度量这种学习的差异性? 作为一名老师, 有责任、有义务对上述问题进行思考和研究。我在计算机语言程序设计课程的教学过程中, 针对以上问题, 对学生的平时作业成绩、期末考试成绩和总成绩进行了统计分析, 用卡方检验定量对照比较了不同班级间的学习差异, 找出了学生学风中存在的问题, 为教学提供了有益的帮助。

所取的样本来自于3个系3个班级共96名学生。总学时46, 其中理论教学26、上机20。课外作业每星期交1次, 课堂练习1次。期末考试为闭卷方式。学习总成绩(以百分制记分)由两部分组成: 平时作业成绩占20%, 期末考试成绩占80%。平时作业采用逐题评分方式, 虽然对教师的要求增高、使教师的工作难度加大, 尤其是批改编写程序的作业费时费力, 但对学生得分的评定却十分公平合理, 且易发现学生学习中的问题。为了防止抄作业现象及保证作业能及时收交上来, 我在开学的第一堂课就宣布了规则: “作业迟交者, 只给60%的作业得分。”也就是说, 倘若别人的作业已发下, 某学生才交作业, 即使全对也只能给其60分, 因为该学生有抄别人作业的嫌疑。对于上机学习, 要求学生事先编好程序, 经上机调试, 得到结果。老师则巡回辅导, 当场解答问题。总之, 在教学过程中, 始终坚持对己严、对学生严, 既鼓励优秀学生发挥潜力, 又不放松对学风差的学生进行过程控制。

1 卡方(χ^2)检验简介

卡方检验是比较两个或两个以上样本率(或比)之间差别的显著性检验。当相比较样本的差别是由本身内在因素引起而非抽样误差所致时, 卡方值就大, 相应的 P 值——反映由抽样误差引起的样本差别

收稿日期: 2001-09-28

作者简介: 徐向阳(1949—), 男, 江苏南京人, 南京工业大学经济管理学院副教授, 硕士生导师。

的概率就小。这时就称两样本差别“有显著性”或“有高度显著性”。反之,卡方值越小, P 值就越大,则称两样本差别“无显著性”。

在实际操作中,可利用列联表模式进行预备计算。设有 R 行 C 列的列联表模式(见表 1):

表 1 列联表模式

组别 \ 水平	1	2	...	C	合计
1	$f_{11}(p_{11})$	$f_{12}(p_{12})$...	$f_{1C}(p_{1C})$	G_1
2	$f_{21}(p_{21})$	$f_{22}(p_{22})$...	$f_{2C}(p_{2C})$	G_2
...
R	$f_{R1}(p_{R1})$	$f_{R2}(p_{R2})$...	$f_{RC}(p_{RC})$	G_R
合计	N_1	N_2	...	N_C	N

表中, f_{RC} —— 第 R 组第 C 个水平的样本频数 ($R=1, 2, \dots; C=1, 2, \dots$)

p_{RC} —— 第 R 组第 C 个水平样本的百分比(率)

G_R —— 第 R 组各个水平样本的频数之和

N_C —— 各组第 C 水平样本的频数之和

卡方值则由下式计算:

$$\chi^2 = N(a_1 + a_2 + \dots + a_k - 1).$$

式中, $a_k = N_C(f_{1C} p_{1C} + f_{2C} p_{2C} + \dots + f_{RC} p_{RC})$ 。

计算出卡方值后,再根据自由度 $n' = (R-1)(C-1)$ 查有关的卡方值表,从而判别各组样本间差别的显著性。当 $\chi^2 \geq \chi_{0.01}^2$ 时, $P \leq 0.01$, 差别有高度显著性; 当 $\chi_{0.01}^2 \geq \chi^2 \geq \chi_{0.05}^2$ 时, $0.01 \leq P \leq 0.05$, 差别有显著性; 当 $\chi^2 < \chi_{0.05}^2$ 时, $P > 0.05$, 差别无显著性。

2 学生成绩的统计分析

根据上述原理,先将各班的平时作业成绩、期终考试成绩和总成绩划分为 5 个档次(水平),分别计算各自的百分比率,汇总在表 2、表 3、表 4 中;然后再进行 3 个班级(组别)的对照比较,计算卡方值;最后根据卡方值进行差异分析,找出教学过程中的某些规律。

表 2 作业成绩统计表

班级 \ 成绩	≥ 90	80 ~ 89	70 ~ 79	60 ~ 69	< 60	合计
安961	1(4.3%)	5(21.7%)	8(34.9%)	7(30.4%)	2(8.7%)	23
管961	6(17.6%)	13(38.3%)	8(23.5%)	5(14.7%)	2(5.9%)	34
商961	8(20.5%)	14(35.9%)	9(23.1%)	3(7.7%)	5(12.8%)	39
合计	15(15.6%)	32(33.3%)	25(26.1%)	15(15.6%)	9(9.4%)	96

表 3 考试成绩统计表

班级 \ 成绩	≥ 90	80 ~ 89	70 ~ 79	60 ~ 69	< 60	合计
安961	7(30.4%)	4(17.4%)	4(17.4%)	6(26.1%)	2(8.7%)	23
管961	2(5.9%)	10(29.4%)	11(32.4%)	5(14.7%)	6(17.6%)	34
商961	3(7.7%)	8(20.5%)	7(18.0%)	5(12.8%)	16(41.0%)	39
合计	12(12.5%)	22(22.9%)	22(22.9%)	16(16.7%)	24(25.0%)	96

表4 总成绩统计表

班级 \ 成绩	≥ 90	80~89	70~79	60~69	< 60	合计
安961	5(21.7%)	5(21.7%)	4(17.4%)	9(39.2%)	0(0.0%)	23
管961	2(5.9%)	11(32.4%)	10(29.4%)	10(29.4%)	1(2.9%)	34
商961	3(7.7%)	10(25.6%)	6(15.4%)	14(35.9%)	6(15.4%)	39
合计	10(10.4%)	26(27.1%)	20(20.8%)	33(34.4%)	7(7.3%)	96

2.1 作业成绩分析

从表2看出,商961班(以下简称商班)作业优秀(≥ 90 分)率最高,占该班人数的20.5%,高于安961班(以下简称安班)和管961班(以下简称管班),尤其高于安班;但同时该班的不及格率也最高,达12.8%。商班和管班的作业优良(≥ 80 分)率均达55%以上,安班的优良率只达26%,说明安班与商班、管班有一定的差异。经两两对比用卡方检验,有如下结果:安—管对比, $\chi^2 = 5.54$; 安—商对比, $\chi^2 = 9.13$ (接近 $\chi^2_{0.05} = 9.49$ 的值); 管—商对比, $\chi^2 = 1.83$ 。它说明了管班、商班学生作业成绩较高,且不分上下,而安班学生作业成绩偏低;但对3个班做总体分析时发现, $\chi^2 = 10.46 < \chi^2_{0.05} = 15.51$,说明了从总体上看,3个班平时作业的差异还达不到显著水平。这表明了(1)平时作业的难度尚不大;(2)存在抄作业现象,尤其表现在管班和商班。

2.2 考试成绩分析

从表3看出,安班考试优秀率最高,达30.4%,不及格率最低,不到10%;而商班优秀率只有7.7%,但不及格率却高达41.0%,实属罕见。管班与商班优秀率差不多,只是不及格率稍好时,达17.6%。经3个班间的卡方检验,得 $\chi^2 = 19.75$,接近 $\chi^2_{0.01} = 20.09$,说明3个班间考试成绩差异基本达高度显著水平。又经两两对比,有下列结果:安—商对比, $\chi^2 = 11.36 > \chi^2_{0.05} = 9.49$,差异显著,安班明显好于商班;安—管对比, $\chi^2 = 8.92$,有一定差异;管—商对比, $\chi^2 = 5.54 < \chi^2_{0.05} = 9.49$,差异不显著。

考试成绩的卡方检验结果表明,试卷难度适中。倘若试题太难、太偏或太简单,都会使卡方值变小而显不出班级间的差异。安班与管班、商班相比,考试优秀率显然高于其他两班,而不及格率却远低于其他两班;尤其与商班相比,差异更为显著。这说明了安班对考试内容掌握较好,复习到位,基本功扎实,学风端正;而商班对考试内容掌握较差,复习不透彻,学风欠端正;管班则介于二者之间。

2.3 总成绩分析

从表4看出,各班的总成绩虽有差距,但与考试成绩的差异相比已大大缩小。3个班级间比较, $\chi^2 = 12.77 < \chi^2_{0.05} = 15.51$,差异无显著意义。这是因为总成绩中作业成绩占20%的缘故。商班作业成绩最好,冲淡了考试成绩最差的负面影响,使不及格率从41%下降到15.4%,同时不及格的绝对人数也大大下降,从16人下降到6人;而安班考试成绩优秀率最高,达30.4%,但由于作业成绩优秀率最低,仅为4.3%,故影响了总成绩的优秀率(21.7%)。这种把作业成绩计入总成绩的做法有两点好处:一是提高了平时作业的价值,促使学生高质量地完成作业,达到及时消化、巩固学习内容的目的;二是冲淡了“一次考试定终身”的影响,以防某些平时不努力的学生到了考试关头突击一下得高分的投机行为,也缓冲了某些平时学习较扎实的学生在考试时发挥欠佳带来的不良结果。当然,按此办法计算总成绩,应以严把作业关、减少作业的抄袭行为相配合,才能使它起到更加有效的作用。

2.4 考试成绩与作业成绩的比较分析

利用表2和表3,可以将各班的考试成绩与作业成绩列在一起进行卡方检验,以找出考试成绩与平时作业成绩之间的某些关联规律。检验结果如下:安班, $\chi^2 = 6.02$,差异不显著;管班, $\chi^2 = 4.86$,差异不显著;商班, $\chi^2 = 10.42 > \chi^2_{0.05} = 9.49$,差异显著。它说明了商班学风欠端正,平时作业虚假成分多,抄作业现象较严重,作业成绩优秀率高并不能代表学习的水平;遇到有一定难度的试卷,就有相当多的人考试不

及格, 显示出考试成绩与作业成绩的严重背离。相反, 安班的学风端正, 作业基本上是自己独立思考完成的, 在复习中能抓住教学的重点内容, 认真理解概念, 善于解剖典型例题, 因此考试成绩与作业成绩的差异并不显著。我在课程结束后, 特意请安班的班主任了解该班考试成绩突出的原因。班主任老师在调查后归纳出以下几点: (1) 学生对计算机语言课程有较浓厚的学习兴趣, 学习主动性较强; (2) 教师讲课重点突出, 尤其对必须掌握的几种基本结构程序讲解细致、分析透彻, 学生接受、消化得好; (3) 学生能围绕教学大纲提出的重点认真复习, 感到试卷中的内容基本上都复习到了; (4) 平时作业基本上是自己独立完成的, 即使作业有错得不到高分, 也明白错在何处, 因此基本知识较扎实。基于上述原因, 安班的考试成绩名列 3 班之首。

3 结 论

利用统计分析中的卡方检验, 可以帮助我们发现教学工作中的某些规律和存在的问题。平时作业成绩的卡方值可以分析各班的作业水平和作业的难度。卡方值越小, 说明作业难度小或者互相抄作业现象较严重; 期末考试成绩的卡方值可以分析出试卷的难度和各班考试的整体水平。试卷难度太大或太小, 都将大大降低卡方值而显示不出考试水平的差异; 作业成绩与考试成绩相比的卡方值可以分析出平时作业与考试的相关性。倘若平时作业质量较高, 能够认真复习迎考的话, 则两者相比的卡方值将下降, 显示出学生的真实学习状态和水平。

Application of χ^2 Test in Analysing Students' Score Difference

XU Xiangyang

(School of Economic Management,

Nanjing University of Technology, Nanjing 210009, China)

Abstract : Applying χ^2 test in the analysis of students' score difference in homework, examinations, and the final evaluation is an effective way to find out the problems in students' study habits and to improve teachers' art of teaching.

Key words : χ^2 test; students' score; analysis of difference

责任编辑 蔡海葆