

浅论相关分析在教育测量中的运用

王成理

(西华大学计算机与数理学院 四川 彭州 611930)

摘要:一门学科要从定性描述进入定量分析,一个重要的步骤就是要对所研究的对象进行测量,这对于物质测量是很方便的,但是教育测量要比物质测量复杂得多,它是不能直接测量的,只有通过出试卷给学生解答才能进行间接测量。而试卷的质量问题是教育测量中一个特别重要的环节。为此,笔者利用相关分析的理论对试卷的质量进行评价。

关键词:教育测量 相关分析

考试是一个重要的环节,是教学结果的反馈,是教育测量的工具。考试的目的不仅仅是为了求得学生阶段性评定值,也不是单纯为了检查教学效果,给老师和学生提供反馈信息。考试不是教学的“终点站”,它是教学的继续。考试更积极的作用在于培养。发现学生的能力、特点和内在潜力等。它主要体现在试卷的编制上,试卷在一定程度上是学习的“指挥棒”。试卷出得好,不仅信度和效度好,而且难度和区分度都能达到理想的要求。如果试卷过难、过易或出一些怪题、偏题;考分偏高或偏低,均不能发挥考试对提高教学质量的积极作用,还会出现“高分低能”、“分数贬值”等不正常现象。这对学习好的学生不能起到鼓励作用,对学习差的学生也起不到督促和鞭策的作用。因此,试卷的制定要有新的指导思想,应该着眼于测量学生的能力,发展学生的“智能”,还要起到预测人才的作用,并能引导学生主动钻研,积极思维,进行创造性的学习。所以,我们必须抓好考试中试卷编制这个关键性的环节。

从事教学的人都有这样的体会,即试卷的编制往往要比备课花更多的精力。一份成功的试卷,必须是能够体现该门课程的较全面的知识,要能够测量出学生的稳定知识水平和学生的智能差异,要能够指引学生自学的途径,同时要能够发现善于创新的学生。因此,一个教师只有在充分认识试卷的编制在考试中的地位和作用,并具有一定的教育测量知识,把试卷的编制当作培养学生能力,发展学生智能的一个重要教学环节,努力使试卷编制科学化,才能进行有效的考试。

试卷编制的基本原则在于保证试卷的质量。从教育测量来看,衡量试卷质量的常用指标有四项:

(一)效度:试卷所要测试的目的达到了如何的程度,是衡量试卷能否准确测出学生在某一学科上真实水平的量。效度可以用测试结果与被测对象在某一学

科上真实水平的相关系数来度量。效度最大值为1,最小值为-1。通常标准化考题的效度系数在0.45以上。具有理想效度的测试,必定具有足够的信度。

(二)信度:一份试卷对同一对象前后测试结果的相关程度,它标志着测试结果的稳定性和可靠性。通常自编考题的信度系数在0.60—0.80之间,标准化考题的信度系数要求在0.90以上,以达到0.95为好。如果一个测试的信度不高,则其效度也不会很高。

(三)难度:衡量试卷各题难易程度的质量指标,反映试题与学生知识水平适合程度的一个量。它不是一个独立的绝对的指标,难度首先应为区分度服务,难度的选择应以提高区分度的数值为出发点。难度的最大值为1,最小值为0。难度系数越接近0.5时,试题的区别力越高。一般说来,在选择试题时,除了尽量使试卷的平均难度接近0.5外,还要使试题的难度适当分散。

(四)区分度:一个试题区分学生水平能力的指标。区分度的高低意味着试题对于能力强与弱的学生在考试分数上区分和鉴别度的高低。一般说来,区分度系数在0.3以上的试题是比较好的;在0.4以上的试题是优良题;而在0.2以下的试题是应该淘汰的。

在教育科学中,教育测量是一个非常重要的环节,而相关分析的理论在教育测量中有着非常广泛的应用,效度、信度、区分度和相关性有着密切关系,它们不同程度的都可以用相关系数来定义和衡量。有了效度、信度、难度和区分度这四种评价指标,我们便可以有效的、全面的对教育进行测量。为此,我特地将一份资料进行相关分析,以便掌握其要领。

一、数据引入

某试卷共有8题,各题的满分分别是5、5、15、15、20、20、10、10分,而根据以往的测试结果,各题的有效分数分别是4、3、12、13、15、17、8、5分,学生考试后对其进行随机抽样,得到资料如下表:

学生	总分	一	二	三	四	五	六	七	八
1	95	5	5	15	15	20	20	8	7
2	91	5	5	15	13	18	17	10	8
3	84	4	5	14	14	15	20	7	5
4	82	5	3	14	15	17	16	6	6
5	73	3	4	13	12	16	14	5	6
6	70	2	3	12	11	16	15	6	5
7	66	3	2	10	13	15	12	7	4
8	61	4	1	10	9	12	18	5	2
9	52	2	4	8	10	10	10	4	4
10	37	1	2	8	7	7	7	3	0

二、相关分析

1、计算试卷的效度:(计算需用下表) $r = \frac{\sum xy - n \bar{x} \bar{y}}{\sqrt{\sum x^2 - n \bar{x}^2} \sqrt{\sum y^2 - n \bar{y}^2}} = 0.954$ (其中 n 表示题目数)

分数	一	二	三	四	五	六	七	八
x	3.4	3.4	11.9	11.9	14.8	14.9	6.1	4.7
y	4	3	12	13	15	17	8	6

2、计算试卷的信度:(计算需用下表)

—	一	二	三	四	五	六	七	八
$\bar{x}$	3.4	3.4	11.9	11.9	14.8	14.9	6.1	4.7
S^2	1.84	1.84	6.69	6.69	10.96	16.29	3.69	5.01

利用系数法计算试卷的信度: $a = \frac{k}{k-1} (1 - \frac{\sum S_i^2}{S^2}) = 0.94$

其中 k 表示题目数,表示第 i 题考试总分的标准差,表示考试总分的标准差。

3、估计试卷各题的难度(H^i) $H^i = 1 - \frac{\bar{x}_i}{W_i}$

其中表示第 i 题的难度,表示所有学生第 i 题的平均分,表示第 i 题的满分。

则, $H^1 = H^2 = 0.32, H^3 = H^4 = 0.21, H^5 = 0.26, H^6 = 0.255, H^7 = 0.39, H^8 = 0.53$ 。

4、利用相关系数法估计试卷各题的区分度

(r_{dj}): $r_{dj} = \frac{\sum x_{ij} y_{ij} - n \bar{x}_i \bar{y}_j}{n S_i S_j}$

其中表示第 r_{dj} 题的区分度, n 为学生人数, x_{ij} 表示第 i 个学生第 j 题的得分,

y_i 表示第 i 个学生的总分, S_i 表示第 j 题的标准差, S 表示总分的标准差。利用下表:

X_{11}	X_{21}	X_{41}	X_{51}	X_{61}	X_{71}	X_{81}	X_{91}	X_{101}	$\bar{X}_1$	S_1	
5	4	5	3	2	3	4	2	1	3.4	1.36	
Y_1	Y_2	Y_3	Y_4	Y_5	Y_6	Y_7	Y_8	Y_9	Y_{10}	$\bar{Y}$	S
95	91	84	82	73	70	66	61	52	37	71.1	17.07

则第 1 题的区分度: $R_{d1} = 0.8665$,同理: $R_{d2} = 0.6721, R_{d3} = 0.9607, R_{d4} = 0.8997$,

$R_{d5} = 0.9419, R_{d6} = 0.8508, R_{d7} = 0.9039$ 。

三、小结:

通过对此试卷的效度、信度、难度、区分度四个方面的分析,可以知道,在教育测量中,一份好的试卷一般要求具有理想的效度,较高的信度,难易程度也较符合标准。试题的难度和区分度虽然是两个不同的指标,但两者之间存在着一定的联系。通常太难或者太易的试题,区分度不高,而难易适中的试题,区分度较高。提高试题的区分度,试卷的效度和信度也相应地提高。试卷质量的高低,不仅直接影

响到考试的效果,也直接影响学生的学习方法和学习成绩评定的客观性。因此,高质量的试卷必须具有充分的代表性,这也是我国教育体制中的一个重要教学环节。

由此可见,相关分析在现代教育测量中占有特别重要的地位,起着举足轻重的作用。

参考文献:

- [1]张弛.应用统计与计算[M].电子科技大学出版社,1998.
- [2]宋乃庆.现代教育统计与测评技术[M].西南师范大学出版社,1998.
- [3]田万海.数学教学测量与评估[M].上海教育出版社,2002.
- [4]张奠宙.数学教育学导论[M].高等教育出版社,2003.
- [5]Gronland.教学中的测验与评价[M].中国轻工业出版社,2003.