

一种离散型随机变量二项分布函数的应用

苏文¹, 包佃清²

(1. 连云港职业技术学院信息工程学院, 江苏 连云港 222006; 2. 连云港师范高等专科学校计算机系, 江苏 连云港 222006)

摘要:简介了经典测试理论,分析了组卷问题的数学模型和常见的组卷算法,提出一种基于二项分布原理的优化随机组卷算法。在组卷过程中,以难度系数为主要控制指标,知识点指标为辅,可以很好地解决试卷的难度控制问题,能够使得试题平均难度系数接近正态分布,符合组卷要求。通过四场模拟考试与数据得出实验数据呈正态分布,信度是83.22,总体来看属于较高。测试的内容效度较高,基本符合预期。

关键词:经典测试理论;二项分布;难度系数

中图分类号:TP301.6

文献标识码:A

测试是教学评价的一个重要手段,传统的卷面测试在时间和空间上受到很大的限制,效率低下。学生作弊、人工阅卷等主观原因,也会影响到测试的公正性。随着测试类型的不断增加及测试标准的不断改进,教育工作者的工作量将会越来越大。传统的测试方式已逐渐不能适应现代测试的要求。目前,随着计算机应用的迅猛发展,网络应用范围的不断扩大,在教育教学中应用计算机已成为一种趋势,从而使得测试的技术手段和媒介也发生了革命性的变化,计算机测试系统随之产生。因此,采用一种合适的组卷算法构建一个操作简单、维护容易、智能性较强的测试系统具有非常重要的意义。

1 经典测试理论

经典测试理论(Classical Test Theory, CTT)又称真分数理论,是一种以测试实得分数为前提条件的理论,被测者在某些方面的真实数值即真分数(True Score)。通过一定测量工具进行测量的值叫观察分数,由于有测量误差存在,所以观察分数并不等于被测者的真实分数,也就是说,观察分数中包含有真分数和误差分数两个因素,其数学模型为:观察分数 = 真分数 + 误差,即:

$$X = T + E \quad (1)$$

这是一个线性模型,它表示在观察分数和真分数之间存

在一种线性关系^[1]。

CTT的评价指标主要有:平均分、难度、区分度、信度和效度^[2]。

2 二项分布函数与组卷算法

2.1 $B(n, p)$ 函数的应用

常见的组卷算法主要有:随机函数法、回溯法和遗传算法几种。随机组卷算法根据状态空间的控制指标,随机地抽取一道试题放入试题库,此过程不断重复,直到组卷完毕或无法从题库中抽取满足条件的试题为止^[3]。该方法结构简单,对于单道题的抽取运行速度较快,但是对于整个组卷过程来说组卷成功率低。本文主要针对随机组卷算法和二项分布函数原理进行组卷优化。

通常在组卷过程中以难度系数为主要控制指标是常用的做法,而区分度、信度和效度三个指标,往往只有以定量的样本,反复实际测试后才有可能获得,不具备可行性。

试卷的难度作为试卷的特征指标,表示组成试卷题目的总体平均难度水平。试卷的难度值可以通过对所有题目的难度进行加权平均的计算方法获得,

$$ND = \frac{\sum_{i=1}^m p_i d_i}{\sum_{i=1}^m p_i} \quad (i = 0, 1, \dots, m) \quad (2)$$

收稿日期:2011-07-30

基金项目:江苏省现代教育技术研究2011年度重点课题(2011-R-19079)

作者简介:苏文(1979-),女,江苏连云港人,讲师,硕士,主要从事计算机图形图像、信息技术、计算机辅助教育方面的研究,(E-mail)anyit@126.com

其中, m 是试卷所含的题目数, p_i 、 d_i 分别是第 i 道题的难度和分数。

随机函数法对于每道试题只有两种可能:选择与放弃。因此,基于随机函数法的抽题概率事件符合离散型随机变量的二项分布函数 $B(n,p)$ [4],

$$p_n(k) = \binom{n}{k} p^k q^{n-k} \quad (3)$$

其中, $k = 0, 1, 2, \dots, n$, n 为正整数, $p > 0, q > 0, p + q = 1$; 二项分布的均值为:

$$Q = n \times p \quad (4)$$

其中, Q 表示试卷的平均难度级别, k 表示难度级别, p^k 表示难度级别为 k 的概率。对于 n, p 固定的二项分布 $B = (n, P)$, 当 n 单调增加时, 概率 $P = \{x = k\}$ 先是单调增加至最大值, 然后单调减少, 两头的概率很小, 可以忽略。这种特性符合一般的组卷标准, 试题的难度系数呈正态分布。

设试卷总分数为 s , 则难度级别为 k 的试题在试卷中的分数为:

$$p(k) = S \times p_n(k) \quad (5)$$

通常利用此模型, 可得各难度级别的试题分数 $p(k)$, 计算步骤如下:

- (1) 确定试题难度的总级别 k , 则 $n = k + 1$ 。
- (2) 根据 n 和 Q 值利用 $Q = n \times p$ 计算出 p 值。
- (3) 将 n 和 p 代入公式(5)计算出 $p_n(1) \sim p_n(k)$ 的值。

(4) 根据 $p_n(1) \sim p_n(k)$ 和 s , 利用公式(5)计算出 $p(1) \sim p(k)$, 如果 $p(1) \sim p(k)$ 之和与 s 有偏差, 则将差值加到与 Q 最接近的难度级别对应的分数中。

采用二项分布函数既能使试题平均难度系数接近理想值, 又能使试题不过于集中在个别难度系数上, 这在很大程度上避免了缺题现象的出现[4]。也能够使得试题平均难度系数接近正态分布, 结合知识点的分布, 比较容易抽出符合标准的试题。

2.2 一种优化的随机组卷算法

组卷问题实质上是一个多目标优化的问题, 即根据一定的约束关系抽取一定试题。用向量坐标表达试题的抽取过程, 记 $Y(V, M, N)$:

$V(v_1, v_2, \dots, v_n)$, 表示一个属性变量集, 即知识点、试题类型以及难度等变量;

$M(m_1, m_2, \dots, m_s)$, 表示一个有限数字集, 有各种属性的取值范围;

$n(n_1, n_2, \dots, n_t)$, 表示一个有限规则集, 表示各试题应满足的约束条件。

试题的抽取过程 Y 是在同时满足 V, M, N 的条件下进行的一个约束满足过程。

本文主要以难度系数为控制指标, 确定试卷难度、题目数量, 依据式(2)~式(5)得到组成试卷所需的每级难度试题数量。然后抽取难度级别为 K 的试题时, 使用

随机函数 $\text{Rand}()$ 生成随机数, 加入知识点比较过滤掉知识点重复的试题, 算法流程图如图 1 所示。

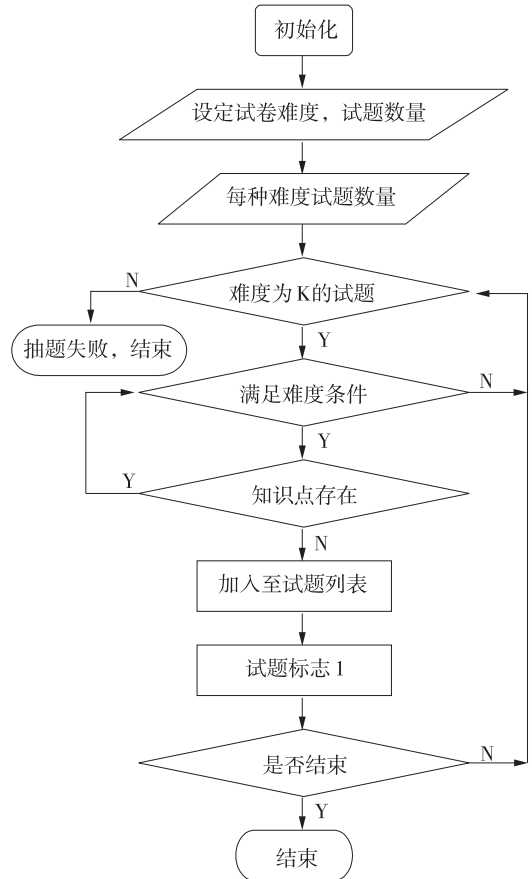


图 1 优化的随机组卷算法流程图

在组卷过程中抽取难度级别为 K 的试题时:

- (1) 根据难度等级、知识点分布、试题类型等参数组成约束条件的总体。
- (2) 选取满足条件的全部试题的非空试题集合, 满足抽题的约束条件: 难度、知识点范围等。
- (3) 用随机函数 $\text{Rand}()$ 从非空试题集合中选择一道试题 K , 并用数组 $a[i]$ 记下这个随机数。
- (4) 重复进行第 3 步操作, 记下随机数 $a[i] + 1$ 。
- (5) 判断 $a[i]$ 是否等于 $a[i] + 1$, 若相等则说明试题抽取重复, 放弃后重新进行随机数产生, 并返回第 3 步。否则读取相关试题。
- (6) 当两者知识点不同时, 把试题加入到试卷中, 并与本试卷中已抽中的其它试题进行知识点比较, 若两者相同, 则转(3), 从集合中再次选题, 并重复本次操作, 否则转(5)。
- (7) 判断是否满足结束条件, 如不满足, 则转(3)重新搜索试题集合, 否则转(8)。
- (8) 最后对生成的试卷中每道试题进行处理, 组卷结束。

3 分析与结论

通过计算机仿真模拟得出实验数据。模拟参加考试的考生有200名,分为4场考试,抽题方式为随机,题库大小为500,每个考生题量为50,难度区间为 $[-2.5 \sim 2.5]$,区分度区间为 $[0.05 \sim 2]$,评价指标涉及平均分、难度、信度^[5]和效度^[6]几个指标,见表1。

表1 实验结果

场数	评价指标		
	平均分	信度	效度
1	62	84.01	0.67
2	64	82.74	0.63
3	61	85.07	0.69
4	65	81.06	0.66

从表1数据可以得出:(1)考试成绩正态分布,说明测验项目适中。在很多考试可以设定为难度偏易,呈现偏态分布曲线以适合大部分考生;(2)在四场考试中判断均相同者占总人数的83.22%,可见测验结果一致的部分占的比重较大,由此说明测验的信度较高;(3)根据克隆巴赫估计效度的方法可以得出结论,四场考试的成

绩相关性系数为0.66,测试的内容效度较高。

在计算机测试中应用离散型随机变量二项分布函数的原理,通过模拟实验及分析得出,实验基本符合预期。

参考文献:

- [1] 陈耿,文登敏.一个基于MIDAS的分布式数据库软件的实现[J].计算机应用,2002,22(9):71-73.
- [2] 包佃清.IRT原理在现代远程教育的应用[J].武汉理工大学学报:信息与管理工程版,2011,33(1):77-82.
- [3] 吉根林.遗传算法研究综述[J].计算机应用与软件,2004,18(2):20-30.
- [4] 林雪明,张钧良,蒋伟钢.基于知识点的试题库组卷算法的建立[J].微机发展,2001,11(2):77-79.
- [5] 袁运开.教育统计、测量与评价[M].北京:中国科学技术出版社,1990.
- [6] 朱德全,宋乃庆.现代教育统计与测评技术[M].重庆:西南师范大学出版社,1999.

Application of Binomial Distribution Function with Discrete Random Variables

SU Wen¹, BAO Dian-qing²

(1. Dept of Information Engineering, Lianyungang Technical College, Lianyungang 222006, China;

2. Dept of Computer Science, Lianyungang Normal College, Lianyungang 222006, China)

Abstract: Classical test theory is firstly introduced, the mathematic model and the common calculation methods of Exam-paper Formation are analyzed, and an optimized random calculation method based on binomial distribution principle is put forward. In the composition of examination paper, difficulty coefficient as the main control index can solve the difficulty of paper control problem, and make the test average difficulty coefficient is close to normal distribution. The conclusion is test scores for the normal distribution, reliability was 83.22 through the simulation experiment. Test validity is high with expectations.

Key words: classical test theory; binomial distribution; degree of difficulty