正态分布函数在试题难度分配算法中的研究

张殿龙
(哈尔滨理工大学 计算中心,黑龙江 哈尔滨 150080)

摘 要:为了客观反映学生对不同难度知识的掌握情况,试卷命题时通常需要按正态分布确定各难度的试题数量.由于数学意义上的正态分布函数是建立在连续空间上的,不能直接适用于试题难度分配算法,因此设计离散数域内正态分布函数值算法就显得更有意义.本文对连续的正态分布函数计算方法进行了改进,解决了试题难度的离散化和试题数量为整型值造成的试题难度正态分分配困难问题.

关键词:正态分布;难度分配;算法
中图分类号: G40 - 057 **文献标志码**: A **文章编号**: 1007- 2683(2009)03- 0013- 03

Research of Normal Distribution Function in the Questions Difficult Allocation Algorithm

ZHANG Dian-long
(Computer Center, Harbin University of Science and Technology, Harbin 150040, China)

Abstract: In order to objectively retlect the grasp of students for knowledge of different difficulty, normal distribution of question numbers of different difficulty is required. From a mathematical view, the normal distribution function is built on a continuous space; it can not be directly applied to allocation algorithm of test questions. Therefore, the numerical algorithm of normal distribution function in continuous space is improved, so that the difficulty of the discretization of question difficulty and the difficulty of normal distribution of question difficulty caused by the integer value of question number can be solved.

Key words: normal; difficult and distribution; the algorithm

1 正态分布函数在试题难度分配算法中的应用现状

对学生成绩的评定是衡量学生个体学习的进展和变化的重要参考因素,是教育者最关心的工作之一[1].常见而有效的评定方法是通过将一定难度和数目的题目组成一张试卷测试学生.难度是试题技术质量的一个重要方面.如果试题难度不当,一个测验就很可能达不到测验目的[2].教育测量学理论认为,对于一般的常模参照测验而言,其目的在于测量个体差异.当试题难度接近 0或接近 1时,即被试在该题上全部答对或全部答错时,该题便无法提供个体差异的任何信息.若每一道题难度值均为 0.5,那么此测验很可能只可区分出好和差两种极端被试的差异,却不能对各种被试作更精细的区分[3].因此在设计试卷时,如何合理地按难度分配试题数量是决定测试效果的重要因素.由于正态曲线有以下几个特点:曲线在期望值处为最高点,以期望值为中心两侧对称,从最高点向两侧缓慢下降,并无限伸

收稿日期: 2008 - 08 - 11
作者简介: 张殿龙 (1965—),男,高级工程师.

© 1994-2009 China Academic Journal Electronic Publishing House. All rights reserved. http://www.cnki.net

延,但始终为正[4].

为了使测验结果能客观地反映被测试者对不同难度的知识掌握情况,试卷命题时通常需要将试题难度按正态分布.各难度的题目被抽中的数量应与正态概率密度函数有正比例关系.即设一套试卷中有 N 道选择题,难度为 X_1 的题目数量为 N_1,占总题数的比例为 Y_1;难度为 X_2 的题目数量为 N_2,占总题数的比例为 Y_2, …;难度为 X_n 的题目数量为 N_n,占总题数的比例为 Y_n.若希望此 N 道选择题难度呈正态分布,则需要满足

$$y = \frac{1}{\sqrt{2\pi}\sigma} e^{-\frac{(x-\mu)^2}{2\sigma^2}}$$

其中:μ为平均难度及期望值;σ为区分度即方差.μ越大,试卷难度越大,整体成绩越低;μ越小,试卷难度越小,整体成绩越高,σ越大,试卷难度越分散,个体成绩的差异区分越明显;σ越小,试卷难度越集中,个体成绩的差异越模糊.在常规考试中,教育者希望通过测验检查学习者对知识的掌握情况,因此宜采用 μ适中而 σ较大的难度分布.在特性选拔考试中,教育者希望通过测试,选择少量带有某种特性的个体学习者,则宜采用 σ较小的正偏态或负偏态的难度分布.

分别取区分度 b为 0.2、0.3、0.4,通过计算机作图显示,如图 1所示.弯曲程度最大的曲线是 $\sigma = 0.2$时的难度分布曲线,弯曲程度最小的曲线是$\sigma = 0.4$时的难度分布曲线,显然,$\sigma < 0.2$时,极难的试题和极简单的试题将不易被随机选中,而 $\sigma > 0.4$时,极难的试题和极简单的试题又太容易被随机选中,因此当 $\sigma \approx 0.3$时,卷面具有较为理想的区分度.

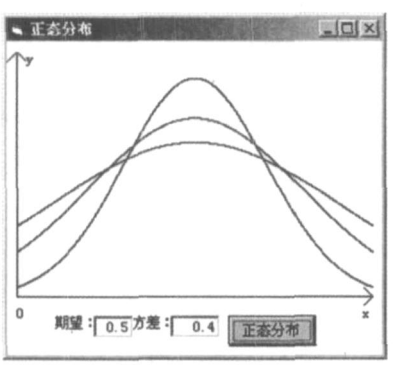

图 1 方差分别为 0.2、0.3、0.4 正态分布曲线

虽然理论上根据难度 X的值,很容易计算出该难度试题数量 Y值,但是正态分布函数的前提是建立在连续函数基础上的,而实际出题过程中,常常由于以下因素不能直接使用正态分布函数计算各难度题目的数量[5].

1)题目难度不能为任意实数.由于受目前技术水平的限制,输入题目,确定题目难度及其他参数仍需要人工参与,因此将难度定为任意实数会在题目录入时耗费大量的时间和人工成本.通常解决的办法是将题目难度划分为若干区间,处于某一难度区间的题目按该区间中值近似表示.建议此类区间个数为 5, 7, 9,区间个数太少将无法将题目的难度区分开,不利于正态分布的近似模拟;区间个数太多则将造成各区间题目数量相对减少,不利于随机抽取,同时增加了录入成本.

2)题目数量不能为任意实数.由于每套题试题数量是有限的,得到的各难度比例 Y需要通过某种折算,还原为整数值.

因此试卷难度的分布不可能达到理论上的正态分布,那么怎样得到比较准确的离散数域内的正态分布难度整数值算法,以及此算法的误差估计,是很有必要研究的[7].

2 近似取整算法

由于理论公式计算在实际试题难度分配中无法实现,这里给出一种将 n道试题按难度 X_i取离散值 {0.1, 0.2, 0.3, 0.4, 0.5, 0.6, 0.7, 0.8, 0.9},并将分配比例按总数采用近似取整的算法.

首先计算出各难度正态分布的理论初值,即

$$Y(i) = \frac{1}{\sqrt{2\pi}\sigma} e^{-\frac{(i/10-\mu)^2}{2\sigma^2}},\ i = 1, 2, 3, 4, 5, 6, 7, 8, 9$$

再将计算理论难度比例

$$P(i) = \frac{Y(i)}{\sum_{i=1}^{9} Y(i)},\ i = 1, 2, 3, 4, 5, 6, 7, 8, 9;$$

然后计算理论难度比例近似整数

$$D(i) = [P(i) \times n + 0.5],\quad i = 1, 2, 3, 4, 5, 6, 7, 8, 9$$

用此算法可很容易得到与理论难度分配比例接近的整数难度分配.这种近似算法在试题总量较大时,比较接近理论值.但是,由于取整计算过程具有舍入误差,有时会产生一些试题总量上的出入.例如题量分别取 10、20、30、60,难度期望取 0.5,区分度方差取 0.3,则各难度试题分配数量如表 1,表明该算法在部分定义域内有误差存在.

© 1994-2009 China Academic Journal Electronic Publishing House. All rights reserved. http://www.cnki.net

表 1　近似取整算法对题量为 10、20、30、60的难度分配结果

题数	难度									总和
	0.1	0.2	0.3	0.4	0.5	0.6	0.7	0.8	0.9	
10	1	1	1	1	2	1	1	1	1	10
20	1	2	2	3	3	3	2	2	1	19
30	2	3	4	4	5	4	4	3	2	31
60	4	6	7	9	9	9	7	6	4	61

3　修正算法

在上面的近似取整算法中,出现了对于个别试题总量,抽取试题总和与要求不符的情况,为了合理地改正这种算法的舍入误差,这里给出一种修正算法.

首先计算近似整数算法得到的题目总数与要求的题目总数的差,即

$$D_{ef} = \sum_{i=1} D(i) - n$$

再根据偏差的数量 E 对各难度题目数量重新调节.如果偏差量 E 为正数,则表明近似题量多于要求的题量,此时为以难度期望为中心,向两侧难度依次减少 1个题目,直到修正后题量与要求的题量一致;如果偏差量 E 为负数,则表明近似题量少于要求的题量,此时为以难度期望为中心,向两侧难度依次增加 1个题目,直到修正后题量与要求的题量一致.

通过修正算法,克服了难度整数分布近似计算中,由于舍入误差造成的试题总量上的误差.题量分别取 10、20、30、60,难度期望取 0.5,区分度方差取 0.3,则各难度试题分配数量如表 2.

表 2　修正算法对题量为 10、20、30、60的难度分配结果

题数	难度									总和
	0.1	0.2	0.3	0.4	0.5	0.6	0.7	0.8	0.9	
10	1	1	1	1	2	1	1	1	1	10
20	1	2	2	3	4	3	2	2	1	20
30	2	3	4	4	4	4	4	3	2	30
60	4	6	7	9	8	9	7	6	4	60

按修正算法期望值取 0.5,方差取 0.3,总题量取 30,得到各难度题目数量分配比例如图 2.从图中可以看出试题难度基本符合正态分布曲线,试题总量满足要求,没有偏差.

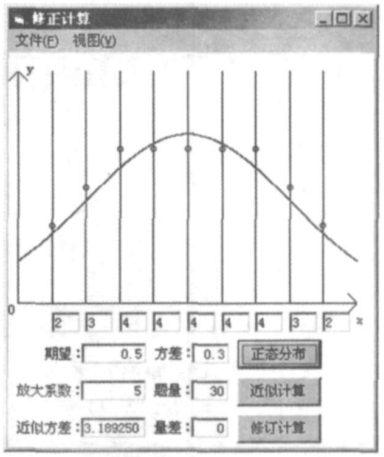

图 2　方差为 0.3 难度为 0.5 题量为 30 的修正后试题难度分布

4　结　语

正态分布概率密度函数是建立在连续实数函数空间的[8],要应用在离散的定义域内取整数解就需要对正态分布密度函数进行定义域的离散化和函数值域的整数化[9],在函数值值于的整数化过程中,可能产生取舍误差[10],需要对简单整数化按需要进行修正.本文介绍了一种正态函数离散化的算法,在试卷命题过程中应用这个算法可以使试卷难度基本符合正态分布.

参 考 文 献:

[1] 陈玉琨.教育评价学[M].北京:人民教育出版社,1999.
[2] 漆书青.略论试题难度[J].江西:江西师范大学学报:哲学社会科学版,1985,2:54-60.
[3] 张敏强.教育测量学[M].北京:人民教育出版社,1998
[4] 王孝玲.教育统计学[M].上海:华东师范大学出版社,2001.
[5] 张国志,鲍　曼,木壮志,等.未知二次损失下正态均值的估计[J].哈尔滨理工大学学报,1998(3):93-97.
[6] 宣仲良.各级难度试题分值比例的确定[J].淮阴师范学院学报:自然科学版,2004(3):24-245.
[7] 陆　明.约束满足问题在题库系统中的应用[J].民营科技,2008(4):44-48.
[8] 徐宝林.分布式通用题库系统研究综述[J].广东白云学院学报,2008,15:38-41.
[9] 朱守业.通用题库数据结构与组卷算法设计[J].现代教育技术,2008(8):101-104.
[10] 刘　文,刘艳伟,刘秋梅.关于试题难度系数调整算法和组卷算法的研究[J].燕山大学学报,2007(3):266-268.

(编辑:付长缨)

© 1994-2009 China Academic Journal Electronic Publishing House. All rights reserved.　http://www.cnki.net