



模糊 S-P 表软件设计及应用

崔爱玉

南京第十三中学 江苏 南京 210008

电邮: caywuli2002@gmail.com

收稿日期: 二零一零年三月二日 (于六月十一日再修定)

内容

- [摘要](#)
 - [引言](#)
 - [模糊 S-P 表的制作](#)
 - [模糊 S-P 表软件的设计流程](#)
 - [结语](#)
 - [参考文献](#)
-

摘要

本文介绍了模糊 S-P 表的制作方法, 利用 C++ 编程语言在 Visual C++ .NET 环境中制作模糊 S-P 表软件。该软件能够准确快速的绘制模糊 S-P 表, 并能为教师和学生提供多种潜在的动态信息, 教师都可在相关数据科学分析的基础上, 调整教学策略。

关键词: 模糊数学; S-P 表; 计算机应用; 大学物理

引言

S-P (Student-Problem) 表是一种学生答题得分表, 它是由日本应义塾大学教授藤田广一先生在 1969 年提出的一种教学评价方法。它是对学生群体、个体和试题等方面进行评价, 是一种适应性很广的方法。本文把模糊数学理论引入到原 S-P 表评价方法中, 突破了原 S-P 表的“0”、“1”记分法的限制, 并且把这种模糊 S-P 表理论应用到软件中, 为更多的教育研究工作者提供了方便, 用户不需要编写复杂的程序, 直接可以进行便利地分析学生的学习动态消息, 因此该软件的研究具备很高的实用价值。



模糊 S-P 表的制作^[1]

所谓 S-P 表, 是把使用选择题型的测验分数变成 1 (答对) 或 0 (答错) 的数据, 并按一定的规则排列的每人每题得分表。排表的规则是: 学生顺序按得分多少由上而下排列, 问题顺序按答对人数的多少由左向右排列, 并画出学生得分曲线, 称 S (student) 线, 画出问题答对人数曲线, 称 P (problem) 线。模糊 S-P 表改进了原 S-P 表中的 0-1 记分法, 原理与其类似。

1.1 把每个题目的得分进行模糊化处理

所谓模糊化处理, 就是将每个学生的每道题的得分除以这道题的满分值, 即该题的答对率, 也称隶属度。如果该学生在该题的得分是满分, 则该题的隶属度就为 1。引入模糊的概念, 可以进一步拓宽 S-P 表仅限于分析选择题的功能。对于一些主观题, 模糊 S-P 表也能加以分析。

1.2 将试题及学号排序

把第 i 个学生的每道题的隶属度相加, 和为 X_i 按照从大到小的顺序自上而下排列, 把第 j 题的每个学生得分的隶属度求和得 Y_j , 由大到小地从左到右地排列, 如图 2 所示。

S	P	序次	1	2	3	4	5	6	7	8	9	10	S线	问题	力学	电学	热学	光学	磁学	原子	量子	相对论	
名次	学号		7	8	2	6	9	4	3	5	10	1	X_i	r_i									
1	9	0.8	0.8	1.0	0.9	0.7	0.9	0.9	0.8	0.5	0.8	8.1	0.2	0.8	0.9	0.9	0.8	0.9	0.8	0.8	0.8	0.7	
2	3	0.9	0.9	0.8	0.8	0.4	0.9	0.8	1.0	0.5	1.0	8.0	0.3	0.8	0.8	0.9	1.0	0.8	0.9	0.9	0.4		
3	8	0.9	0.8	0.6	0.9	0.8	0.7	0.8	0.8	0.6	1.0	7.9	0.3	0.7	0.8	0.7	0.6	0.9	0.9	0.8	0.8		
4	2	0.8	0.9	0.8	0.6	0.6	0.6	0.6	0.8	0.8	0.8	7.7	0.3	0.7	0.6	0.6	0.6	0.6	0.6	0.9	0.8		
5	4	0.8	0.8	1.0	0.6	1.0	0.7	0.7	0.6	0.4	0.8	7.4	0.4	0.7	0.7	0.7	0.6	0.6	0.8	0.8	1.0		
6	1	0.9	0.9	0.8	0.8	0.8	0.6	0.7	0.8	0.9	0.4	7.4	0.4	0.6	0.7	0.6	0.8	0.8	0.9	0.9	0.8		
7	7	0.9	0.9	1.0	0.8	0.7	0.7	0.4	1.0	0.9	0.0	7.4	0.3	0.6	0.4	0.7	1.0	0.8	0.9	0.9	0.7		
8	5	0.7	0.7	0.8	0.4	0.5	0.8	0.9	0.8	0.8	0.8	6.8	0.4	0.7	0.9	0.8	0.8	0.4	0.7	0.7	0.5		
9	10	0.9	0.6	0.8	0.8	0.6	0.7	0.7	0.8	0.5	0.4	6.8	0.3	0.8	0.7	0.7	0.8	0.8	0.9	0.6	0.6		
10	12	1.0	0.6	0.6	0.8	0.9	0.5	0.6	0.4	0.7	0.4	6.7	0.3	0.6	0.6	0.5	0.4	0.8	1.0	0.6	0.9		
11	6	0.9	0.7	0.8	0.8	0.7	0.6	0.6	0.4	0.8	0.2	6.3	0.5	0.5	0.6	0.6	0.4	0.8	0.9	0.7	0.7		
12	11	0.7	0.9	0.6	0.7	0.7	0.6	0.5	0.0	0.6	0.6	5.9	0.4	0.6	0.5	0.6	0.0	0.7	0.7	0.9	0.7		
	Y_j	10.4	9.4	9.2	9.1	8.6	8.5	8.2	8.2	8.0	6.8												

图 2



1.3 画 S 线与 P 线

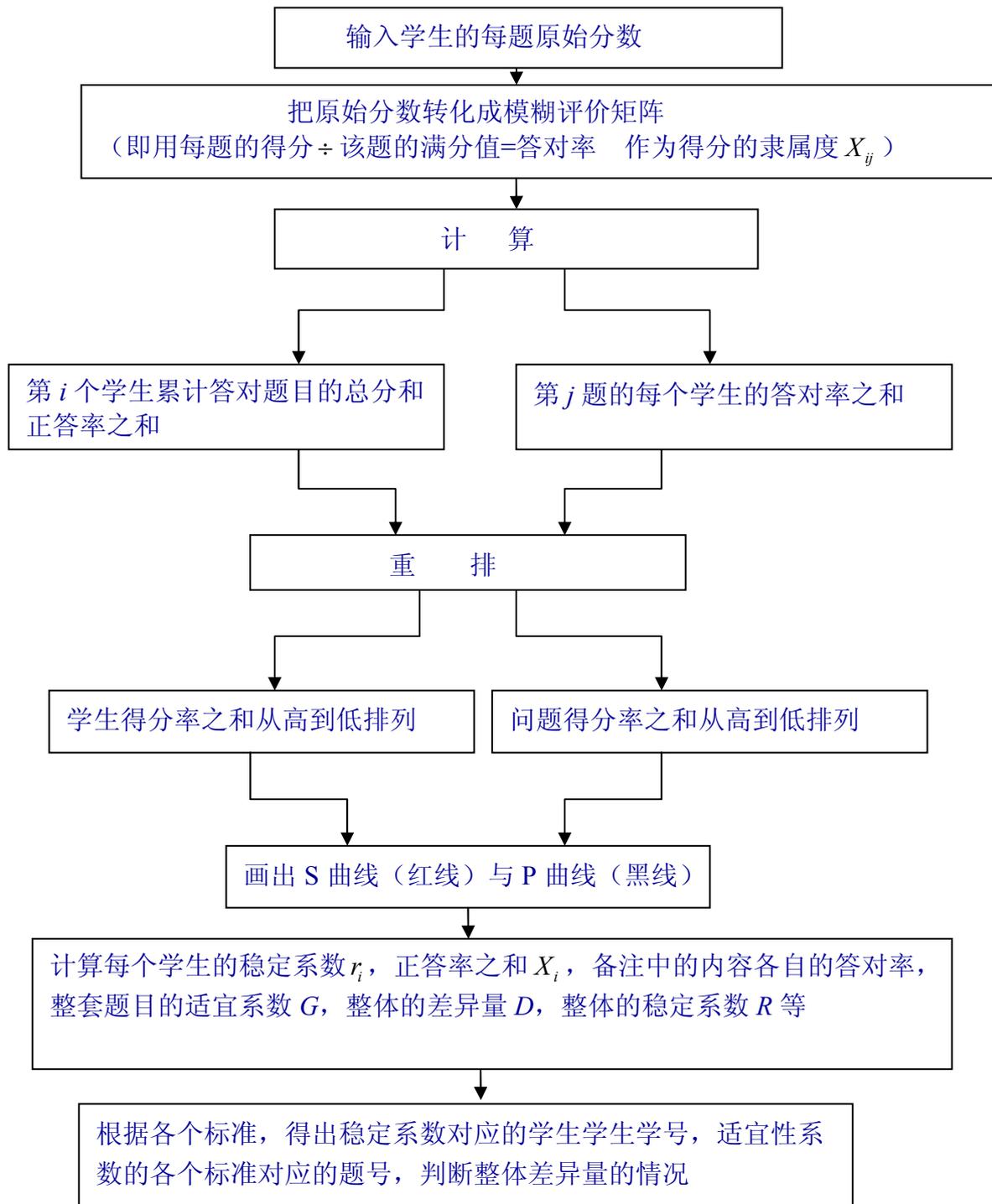
画 S 线：把 X_i 四舍五入，按 X_i 的舍入值数题数题数，例如 9 号同学 $X_9=8.1\approx 8$ ，则在累计 8 道题的右边画一个竖线，学号是 5 号的同学 $X_5=6.8\approx 7$ ，则在第 7 道题的右边画一个竖线，再用横线把不在同一位置的竖线连起来。

画 P 线：同理，把 Y_i 四舍五入，按 Y_i 的舍入值数题数题数，例如第 2 个题目的总隶属度为 $Y_2=9.2\approx 9$ ，于是在累计 9 题的下面画一个横线，以此类推，再用竖线把不在同一行的横线连起来。



模糊 S-P 表软件的设计流程【2】

该软件包括 3 个部分构成：一部分是数据输入说明，一部分是生成模糊 S-P 表，一部分是计算相应的一些回馈信息。整个软件的设计流程图如下：





2.1 输入说明的介绍

输入说明是关于记录每个题目的考察要点的。如果要检查学生对各个知识的掌握程度,以大学物理试卷为例,可以输入“力学”、“热学”、“电学”、“磁学”、“光学”、“原子”、“量子”等部分,参见图 3 的软件截图。这样在图 3 中可以看出每个学生在这个知识点有问题,哪些地方需要进一步加深理解。当然数据说明里的备注部分可以根据用户想要考查的内容而定。比如说,一份大学物理实验试题,在备注中可以标出相应题目所对应考查的哪些能力,如“理解能力”、“动手能力”、“数据处理能力”、“分析设计能力”等。这样在图 2 中 S-P 表的右侧就会出现每位同学的各种能力的答对率,从而可以清楚的了解,每位学生的哪些能力需要进一步提高。

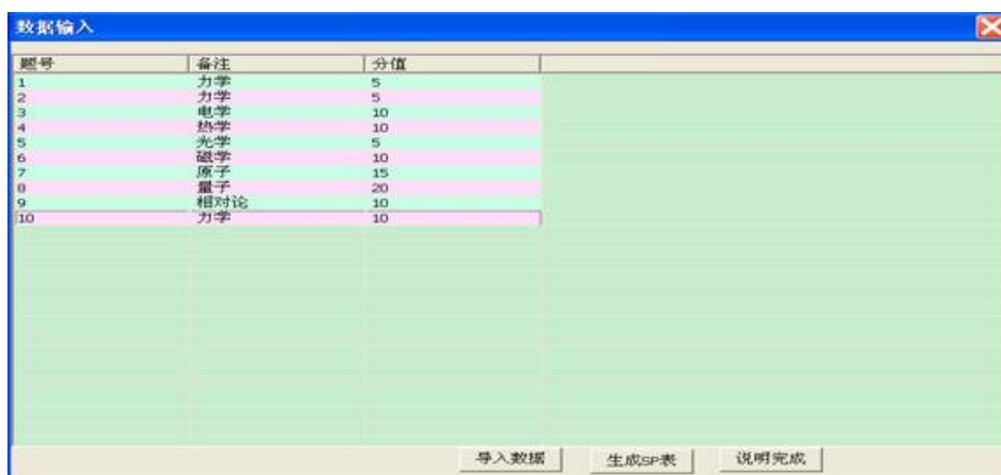


图 3

2.2 生成模糊 S-P 表介绍

本软件在“文件”的下拉列表里点击“新建”时,出现“数据输入”对话框。“导入数据”可以导入在记事本里已经有的考试消息。也可以新建如图 4 所示输入学生数是“12”题目数是“10”,然后再点击“文件”下拉菜单里的“数据说明”,上面已经介绍过了。接着点击“说明完成”,最后点击“生成 S-P 表”。图 2 就是生成后的模糊 S-P 表。



图 4

2.3 回馈信息的介绍

回馈信息主要是计算学生的整体情况, 以及个体评价情况, 还有“知识点回馈”。这些回馈的信息都有相应的公式嵌入在里面, 而相关的系数可以在“参数标准”里输入。根据相应的标准, 可以得出整体评价和个体评价的结果, 如“较好”、“一般”、“较差”。这里介绍一下几个参数的定义。关于每种标准的大小可参见文献[6]。

2.3.1 整体差异量

在 S-P 表中, 整体的差异量是衡量整体应答水平差异的一个指标, 用差异量系数 D (difference index) 来表示, 其计算公式为:

$$D = \frac{\text{s线与p线所围住的数字之和}}{\text{学生数} \times \text{题目数}} = \frac{2(i'j' + \sum_{j=j'+1}^n Y_j - \sum_{i=1}^{i'} X_i)}{m \times n}$$

(注: S 线与 P 线的一个交点为 (i', j') , 学生 $i=1, 2, 3 \dots m$; 题目 $j=1, 2, 3 \dots n$;

原始分数模糊化处理后为 $x_{ij} = \frac{\text{每题的实际得分} b_j}{\text{该题的满分值} a_j} = \text{答对率 (也称隶属度)}$

第 i 个学生对所有题目的正答率之和 $X_i = \sum_{j=1}^n x_{ij}$; 第 j 题目的累计答对率之和 $Y_j = \sum_{i=1}^m x_{ij}$;)



2.3.2 学生群体的稳定系数

稳定性系数是指用来刻画被评价项目的稳定性程度的指标,对于一次考试,是用来评价题目的难易和评卷人的宽严程度的指标,在此有关系如下:

$$R=MF_{左}-MF_{右}$$

($MF_{左}$ 表示 S 线左边学生群体的得分率。 $MF_{右}$ 表示 S 线右边学生群体的得分率)

当稳定性系数 $M=1$ 时, S 线左边全为“1”, 学生状态绝对稳定。经过初步实证分析, 当稳定性系数 $M \geq 0.2$ 时, 测试的稳定性较好; 当 $M \geq 0.4$ 时系统的整体稳定性很高, 学生的被试能力通过测试项目很好地反映了出来。

2.3.3 试题的整体适宜性系数

适宜性系数是指每一个题目能正常区分学生知识水平或能力的指数, 它描述的是试题的适宜性程度. 有如下关系式:

$$G=GF_{上}-GF_{下}$$

($GF_{上}$ 指 P 线上方学生群体的累计得分率, $GF_{下}$ 指 P 线下方学生群体的累计得分率)



图 5

2.3.4 个体稳定系数【4】

通常个体稳定性分析也可以说成个体异常分析.S 线左边、P 在线边不应出现 0, S 线右边、P 线下边不应出现 1, 这时认为学生成绩稳定在各自的水平上。不然则应对学生的学习稳定性设疑。为了考察学生成绩的可靠性, 可用下面的公式计算

$$r_i = \frac{\sum_{j=x'_i+1}^n x_{ij}}{X_i} = \frac{\text{学生 } i \text{ 在 } S \text{ 线右侧数字之和}}{\text{相应的答对率之和}}$$



一般地, $r_i \geq 0.30$, 就要引起注意。 r_i 越大, 则表明学生在测试中表现越反常。

2.3.5 稳定性系数和适宜性系数的关系

稳定性系数和适宜性系数都是反映一次测试的整体水平与质量的指针。稳定性系数反映的是学生应答水平的稳定性程度, 即学生应答水平与测试决策的一致性程度; 而适宜性系数是对测试试题进行整体质量评价。虽然它们评价的角度不同, 但是相对于整个测试项目来说, 它们之间有着密切的联系。当稳定性系数较大时, 学生发挥正常, 应答水平较高, 这很大程度说明试题的质量较高, 设置和编排上较为合理, 此时适宜性系数也会相应较大; 反之, 如果适宜性系数较大时, 试题的得分情况较稳定, 基本上能反映学生的应答水平, 这时学生的应答情况较好, 稳定性系数会比较大。适宜性系数为 1 时, 试题质量很高, 此时学生的反应状态也最好。

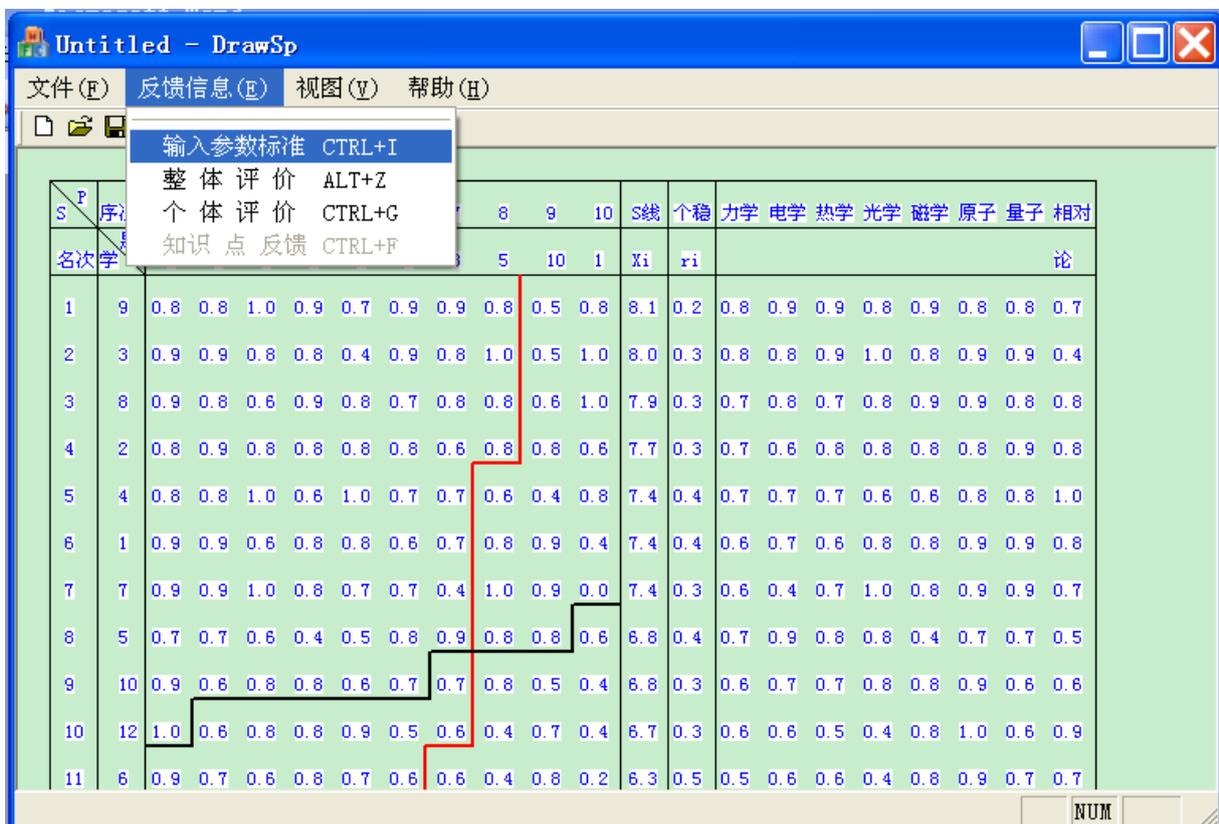


图 6



结语

模糊 S-P 表软件的设计给教师们分析评价学生的学习情况提供了很大的帮助。不需要教师们懂许多编程理论, 直接输入学生的考试信息, 软件会自动画出 S-P 表, 并可以根据 S 线和 P 线的幅度判断出学生差异程度, 软件中还很智能地提供了多种系数的分析。根据各种参数的标准, 会自动计算出各种参数处于什么样的程度, 如“较好”、“一般”、“较差”。总而言之, 本软件的设计对推进教育评价的发展, 具有很强的实用价值。

参考文献

- [1] 张晓晞. 模糊 S-P 表在学生测验结果评价中的应用[J]. 北京联合大学学报: 自然科学版, 2007, 21(2).
- [2] 黄迪生、梁建斌. S-P 表的计算机应用[J]. 新疆师范大学学报: 自然科学版, 2005, 24(1).
- [3] 程志等. F-S-P 表在试卷分析中的应用[J]. 中国高等医学教育, 1996, 6.
- [4] 郑曙东. 《大学物理实验》教学质量的 S-P 表评价方法[J]. 青岛大学学报: 自然科学版, 2004, 12, 22(6).
- [5] 黄焱. 物理 MCAI 与传统教学的 S-P 分析[J]. 昆明师范高等专科学校学报, 2005, 27(4).
- [6] 陈洁. S-P 表评价法的理论与实践[D][硕士学位论文]. 上海: 华东师范大学, 2007.
- [7] 刘新平, 刘存侠. 教育统计与测评导论[M]. 北京: 科学出版社, 2003.