

文章编号: 2095-2163(2020)03-0246-05

中图分类号: TP311.1

文献标志码: A

# 基于 Excel VBA 的试卷质量分析与评价系统

朱西敏<sup>1</sup>, 韩爱庆<sup>2</sup>, 张未未<sup>2</sup>, 王 丽<sup>2</sup>, 唐 燕<sup>2</sup>, 翟 兴<sup>2</sup>

(1 华中科技大学 同济医学院, 武汉 430030; 2 北京中医药大学 管理学院, 北京 100029)

**摘要:** 在试卷评价中引入难度、区分度、信度和效度(四度)等数理统计指标, 设计和开发能全面反映试卷质量的试卷分析系统。该系统一方面可以对试卷指标进行计算和客观分析, 多维评价和发现试卷存在的问题, 辅助提高试卷命题质量; 另一方面可客观反映教师的教学水平与学生的学习效果, 发现教学活动中存在的薄弱环节, 促进教学方法的改革。目前该系统已有近 600 个班次的应用, 已经在本单位各学院, 各专业全面推广, 取得了较好的应用效果。

**关键词:** 试卷; 质量分析; VBA; 评价

## Test paper quality analysis and evaluation system based on Excel VBA

ZHU Ximin<sup>1</sup>, HAN Aiqing<sup>2</sup>, ZHANG Weiwei<sup>2</sup>, WANG Li<sup>2</sup>, TANG Yan<sup>2</sup>, ZHAI Xing<sup>2</sup>

(1 Tongji Medical College of Huazhong University of Science & Technology, Wuhan 430030, China;

2 School of Management, Beijing University of Chinese Medicine, Beijing 100029, China)

**[Abstract]** In the evaluation of test papers, mathematical statistics such as difficulty, discrimination, reliability and validity (four degrees) are introduced, and the test paper analysis system which can fully reflect the quality of test papers is designed and developed. On the one hand, the system can calculate and objectively analyze the four degrees of the test paper, multi-dimensionally evaluate and discover the problems existing in the test paper to help improve the quality of the test paper proposition; on the other hand, it can objectively reflect the teaching level of the teacher and the learning effect of the students, discover the weak links in the teaching activities, promote the reform of teaching methods. At present, the system has nearly 600 shifts of application, and has been widely promoted in various colleges of the author's university and has achieved good application results.

**[Key words]** test paper; quality analysis; VBA; evaluation

## 0 引言

试卷分析是教学工作的重要组成部分, 通过有效的试卷分析, 一方面可以对试卷本身进行合理评价, 提高教师的命题水平; 另一方面可评价学生的能力, 判断试卷的整体价值, 促进教学方法的改进和改革。查阅相关文献, 已有高校引进试卷分析系统, 以健全试卷质量评价体系, 不同的高校也形成了不同的基于多个评价指标的分析理论。基于中医药院校特殊的学科性质, 本文通过各类型题目(如选择题、填空题、判断题)的答题情况的整体分析, 借助数理统计方法, 引入对试卷难度、区分度、信度和效度等深入、专业的分析。对此拟展开研究论述如下。

## 1 问题描述

目前, 大部分教务教学系统可将期末成绩自动生成试卷分析表, 对期末成绩进行初步分析。显然, 当前的试卷分析提供的评价信息严重不足。一方面, 该试卷分析结果是基于总成绩的分析, 无法体现

细节, 未全面、合理地反映试题信息, 对有针对性的修订试卷, 提高教师的命题水平并不具备相应借鉴意义; 另一方面未反映出试卷的难度、区分度和信度, 无法更好地进行教学设计, 改进教学方法, 提升教学质量。所以无法从该分析获知试卷评估及改进试卷质量所需的具体信息。本次研究中, 将用到的学生成绩分布图如图 1 所示。

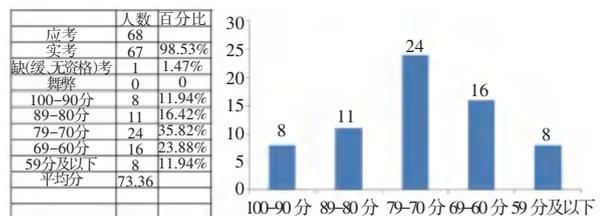


图 1 学生成绩分布图

Fig. 1 Score distribution

## 2 评价指标研究

### 2.1 难度

难度是指试题的难易程度, 反映被试对象对该

基金项目: 北京中医药大学教育教学课题(XJZX1804)。

作者简介: 朱西敏(1997-), 女, 硕士研究生, 主要研究方向: 医学数据分析; 韩爱庆(1979-), 男, 副教授, 主要研究方向: 数据库技术、医学数据分析、人工智能。

通讯作者: 韩爱庆 Email: aqhan@hotmail.com

收稿日期: 2019-12-11

题的得分率,是评估此次考试的一个重要指标。一道题目,如得分率高则难度小,否则难度就大。难度的计算公式为:

$$P = X/W. \quad (1)$$

其中,  $P$  为难度;  $X$  为样本平均得分;  $W$  为试卷总分。

一般将难度值大于和等于 0.7 的试题定为容易题;大于 0.4 和小于 0.7 的定为中档题;小于和等于 0.4 的试题定为难题。如满分 100 分的试题,考生平均得分 95 分,则难度为  $95/100=0.95$ ,该试卷较容易。

## 2.2 区分度

区分度是指试题对被测对象实际水平的鉴别能力,是把考生区别开来的统计量。区分度高的试题能将不同水平的考生区分开来,水平高的考生得高分,水平低的考生得低分。区分度高的考试,优秀、一般、差三个层次的学生都有一定比例,如果某一分数区间学生相对集中,高分太多或不及格太多的考试,则区分度低。区分度计算公式为:

$$D = 2(XH - XL)/W. \quad (2)$$

其中,  $D$  为试卷区分度;  $XH$  为 27% 高分组平均分;  $XL$  为 27% 低分组平均分;  $W$  为试卷总分。

一般认为,试题的区分度在 0.4 以上表明此题的区分度很好,0.3~0.39 表明此题的区分度较好,0.2~0.29 表明此题的区分度不太好需要修改,0.19 以下表明此题的区分度不好应淘汰。如,一份满分 100 分的试卷,高分组平均得分 90 分,低分组平均得分 60 分,则区分度为  $2 \times (90-60)/100=0.6$ ,该试卷区分度较好。

## 2.3 信度

信度是指使用同一试卷对考生重复测验时,或使用 2 个平行试卷对考生测验时,所得测验分数的一致性和稳定性程度。信度除了可检验成绩的真实性外,还要说明题目涉及的内容与教学的相关程度。信度的计算常采用 Cronbach 公式,即:

$$r = \frac{k}{k-1} \frac{\sum_{j=1}^k s_j^2 - \frac{(\sum_{j=1}^k s_j)^2}{k}}{s^2} \quad (3)$$

其中,  $r$  为试卷信度;  $s^2$  为全卷得分的方差;  $s_j^2$  为第  $j$  题的方差;  $k$  为试题数。

一般认为信度值大于等于 0.9,表明此题的一致性非常好;在 0.8~0.9 之间表明此题一致性好;0.7~0.8 之间表明此题的一致性可以接受;在 0.6~0.7 之间表明此题的一致性需较大修订;信度值小于 0.6

表明此题的一致性不合格。

## 2.4 效度

效度是指测量结果的正确结果,即测量结果与测量的目标之间的接近程度。效度可以从多个角度进行分析解释,即:内容效度、结构效度、表面效度等。其中,在试题研究中,内容效度是比较重要的,即根据测验目的确定的测验内容所引起的预期反应达到测验目的的程度。本研究通过对各试题之间做线性 pearson 相关分析,得到的相关系数值即效度值。测量结果与目标越吻合,效度值越高;反之,效度值越低。对应的数学公式可写为:

$$r = \frac{\sum (X - \bar{X})(Y - \bar{Y})}{\sqrt{\sum (X - \bar{X})^2 \sum (Y - \bar{Y})^2}}. \quad (4)$$

其中,  $r$  为 pearson 相关系数,  $X, Y$  分别为比较的试题的分数值。

一般认为,试题的效度值在 0.7 以上表明此题的效度好;0.4~0.7 表明此题的效度较好;0.4 以下表明此题的效度较差,试题与所预测的考核目标不符。

## 3 结构、界面及相关设置

系统的总体功能模块如图 2 所示。由图 2 可知,该系统共包括数据录入、数据校验、指标分析、生成分析报告等多个管理模块。对此可做阐释分述如下。

(1) 分数录入。教师按照题型分别录入每一位同学在该题型的得分,同时录入该题型的满分值。

(2) 数据校验。点击“数据检查”按钮,即可检查数据的输入是否符合逻辑,自动标识出类型错误或空缺值等非法数据。

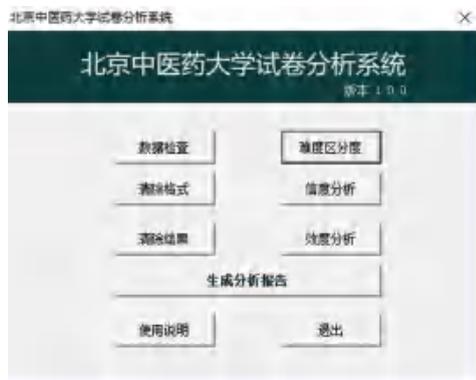


图 2 主界面

Fig. 2 Main interface

(3) 难度区分度分析。点击“难度区分度”按钮,计算并输出平均分、标准差、难度值、难度说明、区分度值和区分效果等分析结果,如图 3 所示。

项目	平均分	标准差	难度值	难度说明	区分度值	区分效果
选择题	22.17	3.06	0.74	中档	0.45	非常好
填空题	6.62	0.85	0.96	容易	0.3	好
名词解释题	11.79	1.91	0.84	容易	0.37	非常好
判断题	6.08	1.18	0.81	容易	0.52	非常好
简答题	8.66	2.65	0.56	难	0.95	非常好
论述题	11.75	1.36	0.85	容易	0.34	好
总分	78.08	5.45	0.78	容易		

图3 难度区分度结果图

Fig. 3 The result of difficulty and discrimination

(4) 信度分析与效度分析。分别单击信度分析、效度分析按钮,进行信度和效度分析,结果如图4、图5所示。

图4 信度分析结果图

Fig. 4 Trust level analysis

	选择题	填空题	名词解释题	判断题	简答题	论述题	总分
选择题	1						
填空题	0.363	1					
名词解释题	0.451	0.18	1				
判断题	0.257	0.1	0.143	1			
简答题	0.293	0.071	0.17	0.17	1		
论述题	0.095	0.185	-0.054	0.331	0.403	1	
总分	0.767	0.391	0.578	0.46	0.766	0.46	1

图5 效度分析结果图

Fig. 5 Validity analysis

(5) 生成 word 格式的分析报告。点击“生成分析报告”按钮。截取部分报告内容见图6。

## 北京中医药大学 试卷分析报告

教务处

难度区分度分析						
项目	平均分	标准差	难度值	难度说明	区分度值	区分效果
1. 选择题	24.4	4.35	0.69	中档	0.61	非常好
2. 填空	6.69	2.05	0.87	中档	1	非常好
3. 名词解释题	7.2	2.06	0.72	中档	0.97	非常好
4. 简答题	8.84	1.77	0.88	容易	0.71	非常好
5. 论述题	7.56	2.51	0.76	容易	1.19	非常好
6. 试卷分析题	5.61	3.25	0.55	难	1.54	非常好
总分	68.57	12.86	0.7	中档	0.64	非常好

图6 分析报告

Fig. 6 Analysis report

## 4 核心算法及其实现

### 4.1 信度算法及实现

#### 4.1.1 信度算法

信度算法的基本原理是:对于一个好的问卷,单个题目(考察某一主题)的组内成绩变异一般比较小,而不同题目间(组间)的成绩变异一般比较大,组内变异与组间变异之和得到总变异,组内变异在总变异中占比越小,试卷的质量越好。算法如下:

**步骤 1** 计算各个题型的方差,即组内变异。

**步骤 2** 计算每个样本的方差,即计算每个样本在各个题型和总分上的方差。

**步骤 3** 计算每个样本方差的和,即总方差。

**步骤 4** 计算自由度,因包含总分列,所以自由度=题型数量+1。

**步骤 5** 将总变异、组内变异、自由度代入 Cronbach 公式,计算信度值。

**步骤 6** 按照以上方法,逐一计算删除各个题型后的信度值。

#### 4.1.2 信度算法实现

算法实现中,首先在主程序中复制 data 表中的原始数据到辅助表 data\_rel 表,然后调用以下 rel\_val() 函数。在以下函数中,利用辅助表分别计算各个题型的变异以及总变异,再将中间结果分别代入 Cronbach 公式,计算信度。若要计算删除某个题型后的试卷信度值,只需利用循环方法逐一删除 data\_rel 表中各个题型,并循环调用 rel\_val 即可。程序代码具体如下:

Function rel\_val() As Single

Dim ws1 As Worksheet, ws2 As Worksheet, ws3 As Worksheet

Dim ii as integer, jj as integer

Dim rel As Single, rng1 As Range, rng2 As Range

Set ws1 = Worksheets ("Data")

Set ws2 = Worksheets ("Data\_rel")

注释:ws1 指向 data 表,ws2 指向辅助表 data\_rel 表,辅助计算行列方差

ws1.Range ("a1").CurrentRegion.Copy ws2.Range ("a1")

注释:复制 data 表数据到 data\_rel 表

ii = ws2.Cells(65536, 1).End(xlUp).Row

注释:ii 分别为 data\_rel 表的行数

jj = ws2.Cells(1, 1).End(xlToRight).Column

注释:jj 分别为 data\_rel 表的列数

For j = 2 To jj

注释:for 循环用于计算每种题型的列方差,计算结果存储到该列下方

Set rng1 = Range(ws2.Cells(2, j), ws2.Cells(ii, j))

ws2.Cells(ii + 1, j) = (Application.WorksheetFunction.StDev(rng1))^2

Next

For i = 2 To ii

注释:for 循环计算每一行(每个人)单项成绩

和总成绩的和

```
Set rng2 = Range(ws2.Cells(i, 2), ws2.
Cells(i, jj))
ws2.Cells(i, jj + 1) = (Application.
WorksheetFunction.Sum(rng2))
```

Next

注释: rng1, rng2 分别指向末行,末列,即,各个题型方差,和每个人的成绩和

```
Set rng1 = Range(ws2.Cells(ii + 1, 2), ws2.
Cells(ii + 1, jj))
```

```
Set rng2 = Range(ws2.Cells(2, jj + 1), ws2.
Cells(ii, jj + 1))
```

```
ssi2 = Application.WorksheetFunction.Sum(rng1)
```

注释:计算各题型成绩方差的和

```
s2 = (Application.WorksheetFunction.
StDev(rng2)) ^ 2
```

注释:计算每个人总成绩的方差

```
k = jj - 1
```

```
rel = (k / (k - 1)) * (1 - ssi2 / s2)
```

注释:Cronbach 公式,计算信度

```
rel_val = rel
```

注释:返回计算结果

End Function

#### 4.2 效度算法及实现

效度计算本质上是计算各试题之间的 pearson 线性相关系数,在 Excel VBA 中,可直接调用 Application.Correl(array1, array2) 返回 2 个向量的相关系数,例如,可以将选择题(*i* 列)和填空题得分(*j* 列)分别作为实参传递给 Correl 函数,计算两列之间的相关系数。

本研究基于 Correl 函数设计了专门的 MCorrelation 函数,对全部题目分别做两两相关分析,研发代码及代码解释如下。

```
Function MCorrelation (rng As Range)
```

注释: rng 为各人各题目得分矩阵

```
Dim x As Variant, y As Variant
```

```
Dim s As Integer, t As Integer, c() As Variant
```

```
cols = rng.Columns.Count
```

```
ReDim c (cols, cols)
```

注释:cols 为列数,矩阵 c() 用于存储效度值

```
For i = 1 To cols Step 1
```

```
For j = 1 To i Step 1
```

```
c(i, j) = Application.Correl(Application.
```

```
Index(rng, , i), Application.Index(rng, , j))
```

注释:Correl 函数功能是计算效度

```
Next j
```

```
Next i
```

```
MCorrelation = c
```

```
End Function
```

#### 4.3 生成分析报告

利用 Excel VBA 计算生成的难度、区分度、信度和效度等各项分析结果分散于同一工作簿的不同的表中。基于应用需求,需将各项分析指标整合为一个系统的 Word 版分析报告,本系统借助 Excel VBA 实现了该功能。

Excel VBA 具备创建 Word 对象,并对 Word 对象进行格式控制,添加、修改及删除文本和图表等功能。研究中得到的创建 Word 对象并将信度分析结果输出到分析报告的代码具体如下。

```
Public Sub OutputAnalysisReport()
```

注:本代码展示新建 Word 对象和输出信度分析结果

```
Dim appWD As Object
```

```
Dim strWd As String
```

```
Dim ws As Worksheet
```

```
Set appWD = CreateObject("Word.Application")
```

```
appWD.Documents.Add
```

注:创建 Word 对象并新建一个空文档

```
appWD.Selection.Font.Size = 14
```

```
appWD.Selection.Font.Name = "宋体"
```

```
appWD.Selection.Font.Bold = True
```

注:设置字体格式

```
strWd = "二、信度分析"
```

```
appWD.Selection.TypeParagraph
```

```
appWD.Selection.TypeText (strWd)
```

注:向 Word 输入标题“二、信度分析”

```
Set ws = Worksheets("信度分析")
```

```
appWD.Selection.TypeParagraph
```

```
ws.Range("a1").CurrentRegion.Copy
```

```
With appWD
```

```
.Selection.PasteExcelTable False, False, False
```

```
End With
```

注:将“信度分析”表中的信度分析结果输出到 Word 中

```
appWD.Selection.TypeParagraph
```

```
Set ws = Worksheets("说明")
```

```
ws.Shapes("pic2").Copy
```

```
appWD.Selection.Paste (下转第 255 页)
```