基于学生成绩分析的教师命题科学性评价[[1]](#footnote-0)

——数据挖掘与特征描述的应用

陈喜华1，黄海宁²，黄沛杰³

1. **广州工商学院教务处，广东 广州510850；**

**2．广州工商学院外语系，广东 广州510850；**

**3.华南农业大学数学与信息学院，广东 广州510642）**

**摘要：**使用Excel函数和特征描述数据分析方法，并借助SPSS Statistics软件，分析学生成绩的离散程度、区分度、难易程度、总体水平特征分布与频率分布，以便评价教师命题的科学性，帮助教师准确了解与掌握学生的成绩分布及对某一门课程的掌握与理解程度，分析结果也能辅助教师科学命题，提高试题的编制质量与教学质量。

**关键词：**数据挖掘；特征描述；SPSS Statistics；命题质量；科学性分析

**中图分类号：**TP399 **文献标识码：A 文章编号：**

学生成绩作为课程考核的最终结果，从客观上反映了学生对某一门课程知识的掌握程度，也是评价教学质量好坏与试题质量高低的重要标准。如何通过这些成绩数据，挖掘隐含在学生成绩中有用的规则与模式，用于评价教师命题的科学性，促进教师改进教学方法，辅助科学命题，提高试卷的编制质量，从而更加有效地反馈教学信息。数据挖掘技术便能有效解决这一难题。

**一、数据挖掘概述**

数据挖掘（Data Mining）就是从大量的、不完全的、有噪声的、模糊的、随机的数据中提取隐含在其中的、人们事先不知道的、但又潜在有用的信息与知识的过程[1]。数据挖掘主要用于决策支持、预测预警、偏差侦测、查询优化、异常分析等。它是一门综合了统计学、数据库、人工智能、可视化多个领域的交叉学科。根据不同挖掘任务，可以将其分为描述和预测两类。数据挖掘的方法有：特征描述、聚类分析、关联分析、分类和预测、离群点分析、演变分析等。本文使用特征描述的方法来分析学生成绩。特征描述即数据总结，就是用简练、精准的方式描述数据的概念，并概括其整体特征，实现对原始数据的总体把握[2]。

数据挖掘是一个反复进行、不断优化的过程，如图1所示，需要在中不断分析、调整、再执行。因此，在开展数据挖掘之前，要先确定目标，选择合适的数据挖掘方法，再根据选择的方法确定挖掘步骤，才能保证数据挖掘按部就班进行并取得预期的结果。其主要步骤有四步：一是确定业务对象。根据业务对象的特征描述，明确需求分析与挖掘目标，清晰地定义业务问题，再预测可能存在的风险和结果。二是数据准备，即数据预处理。原始数据结构复杂，数据量大，且存在冗余、空缺、带噪声等现象，容易影响挖掘结果。因此，需要先对数据进行预处理，否则，很容易对结果的准确度产生较大影响。数据预处理可分为数据选取、数据集成与清理和数据转换三个步骤。三是数据挖掘。按照第一步确定的需求分析与挖掘目标，选择合适的挖掘算法与工具，然后建立模型，并进行挖掘分析。四是解释和评价。分析与验证挖掘结果是否合理、有效，并将得到的有用的规则与模式提供给决策者。

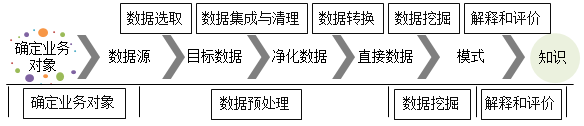


图1数据挖掘过程

**二、面向教师的命题科学性评价****体系及指标**

学生成绩的高低，往往跟教师命题是否科学合理存在一定的联系，因此，提高教师命题的科学性就显得尤为重要。命题是教育测量的重要步骤，决定测试的成功与否。它能有效检验学生的学习效果以及教师的教学效果。对学生来说，能有效检验其学习效果的试题必定具有较好的信度和效度。对教师来说，命题直接影响着教与学的侧重点，科学合理的命题能有效区分学生的能力与层次，方便教师进行教学改革。作为教师，命题应注重考查学生能力，激发学生内驱力，确保命题的科学性与合理性。本文通过分析学生成绩，侧面评价教师命题的科学性。具体评价指标如图2所示。

图2教师命题科学性评价指标

**（一）简单指标**

通过学生成绩来评价教师命题的科学性，往往使用简单的数据统计分析方法，因此，得到的也是评价命题科学性的简单指标，这些指标包括利用Excel软件求出学生成绩的总分、最高分、最低分、平均分等。

平均分（AVERAGE,MEAN）就是求出给定学生成绩的算术平均值。平均分从整体上反映了学生对某一课程知识的掌握程度和理解水平。其计算公式如下：

，n为学生人数。

根据教学实践，在某一次考试中，假如满分为100分，如果，属于不正常现象，表明试题难度太大或教师的教学方法有问题；如果，表明试题较难或学生整体水平较低；如果，表明试题较简单或学生整体水平较高；如果，则表明试题难度适中，比较符合学生的实际水平[3]。使用平均分不仅简单，而且能客观反映学生成绩的总体水平，所以，具有很大的参考价值。

**（二）特征描述指标**

为了更好地评价教师命题的科学性，还需要利用数据挖掘的特征描述方法来反映学生成绩的分布规律，如描述和度量给定学生成绩的百分位值、集中趋势、离散程度和分布形态。在统计学中，通常借助SPSS Statistics软件对数据做进一步描述和度量。

1.百分位值

反映一组数据的集中程度通常使用四分位数（Quartiles）、分割点（Cut Points）及百分位数（Percentile）来表示。在统计学中，把所有数值由小到大排列后平均分成四等份，处在三个分割点位置的数值就是四分位数。如果将一组数据从小到大进行排序，并计算相应的累计百分位，则某一百分位所对应数据的值就称为这一百分位的百分位数[4]。

2.集中趋势

集中趋势表示被观察值向某一中心值靠拢的程度，它反映了数据所聚集的位置。除了可以用平均分、总分来表示，也可以使用中位数和众数来表示。把一组数据从小到大排序，正中间的一个就是中位数（Median）。如果数据有偶数个，则取最中间两个数的平均值作为中位数[5]。众数（Mode）是在一组数据中，出现次数最多的数据。

3.离散程度

离散程度描述的是各个数据之间差异的大小，通常用方差或标准（偏）差来度量。方差（Variance，用表示）用来度量随机变量与其数学期望（即均值）之间的偏离程度。标准差（Standard Deviation，用表示）用来反映学生成绩的离散程度。标准差是方差的算术平方根。，表明大部分学生成绩与平均分差异比较明显，也就是分布比较离散；相反，，表明大部分学生成绩比较接近平均分，也就是分布比较集中。在分析学生成绩时，如果发现标准差太大，表明学生成绩比较分散，差距明显，表明可能存在试题难度过大或教师在教学时缺少考虑接受能力中低程度的学生，导致他们考试不理想。标准差的计算公式为：

，不过这是特殊情况，当所分析的学生成绩很多时，出现这种概率的可能性几乎为0。。

在实际应用中，方差（标准差）或平均值并不适用比较多组数据的离散程度，此时，可以用标准差系数来衡量。标准差系数又称均方差系数，是一个用来描述和度量观察值离散程度的相对指标，可用来比较分析不同年级、不同专业、不同课程的学生成绩。其计算公式如下：

，表明学生成绩越分散，其水平差距就越大；，表明学生成绩越集中，其水平差距就越小。一般地，的值在5%~35%之间比较合理。如果试题难度适中，，这样，每位学生就能同步发展。

4.分布形态

根据统计学原理，通过对学生成绩进行统计分析，发现学生的学习能力具有一定的规律性，而且符合正态分布曲线，如图3所示，正常的学生成绩也应符合此规律。通常，可以用“中间高、两头低”来衡量学生成绩符合正态分布的程度，即大多数学生的成绩为中等或良好，少部分学生的成绩为优秀或不及格。完美的正态分布不太可能出现，即使教师的教学行为一致，学生的学习效果也不可能一样。通常可以用偏度和峰度来定量描述试题的区分度、难易程度或教师的教学方法是否适合多数学生。偏度（Skewness，用*bs*表示）用来度量正态分布的偏斜方向及程度，峰度（Kurtosis，用*bk*表示）反映了峰部的尖度，用来定量描述正态分布的陡峭程度。

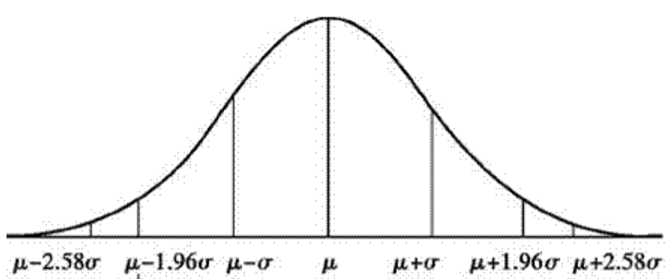


图3正态分布曲线

如果曲线向右偏移（正偏态），即*bs*>0，说明高分的学生人数较少，低分的学生人数偏多，说明试题难度偏大或教师的教学方法不能被多数学生认可；如果曲线向左偏移（负偏态），即*bs*<0，情况刚好相反，说明试题难度较低或多数学生能够认可教师的教学方法；如果曲线基本在中间（正态分布），即*bs*≈0，可以认为是对称的，也就是说试题难度与学生的学习能力相适应或教师在命题时能够较好地把握学生的学习状态。

曲线的中间部分代表大部分学生的成绩，在正态分布情况下，*bk*=3，表明试题难度比例较为合理；*bk*>3，表明曲线较陡峭，学生成绩比较集中，暗示着试题的区分度不够明显或教师的教学方法适合多数学生；*bk*<3，表明曲线较扁平，学生成绩比较离散，暗示着试题的区分度比较明显或教师在教学过程中可能没有充分考虑到多数学生对知识的掌握程度，因此，学生成绩差异较大。值得注意的是，在SPSS等软件中，为了方便比较，默认已将峰度减去3，因此，正态分布的峰度是0。

**（三）频率分布指标**

对学生成绩进行统计分析，通常还需要借助可视化图表（如直方图），以便从学生成绩的总体情况直观地判断教师命题的科学性。如果将成绩划分为无穷个区间时，理论上直方图也可以表示学生成绩的实际分布，将直方图顶端中间的点连成线，便是一条近乎光滑的曲线，这条曲线也就是学习成绩的分布曲线。

**三、基于数据挖掘和特征描述的学生成绩分析**

为了验证教师命题的科学性，本文选取广州GS学院2015级全体参加全国高等学校计算机水平考试的本科学生成绩为研究对象，通过排除缺考与个别成绩为0分的异常数据后，得到3006个有效成绩，再借助SPSS Statistics 22软件来描述和度量本次考试的学生成绩，并评估此次考试的教师命题的科学性和教师的教学效果。从该校教务系统导出要分析的学生成绩，导出默认格式是Excel文件，经过简单处理后，导入到SPSS Statistics中。经过数据挖掘，对学生成绩的总体特征描述如下：

**（一）简单分析**

根据学生成绩统计出学生成绩的个数、平均值、最高分、最低分和总分。依次选择“Analyze—DescriptiveStatistics—Frequencies”，将“成绩”移到Variable(s)中，再选择“Statistics”按钮，勾选“Mean”“Sum”“Minimum”“Maximum”, 其他按默认设置。切换到SPSS StatisticsViewer，便可查看分析结果，如图4所示。

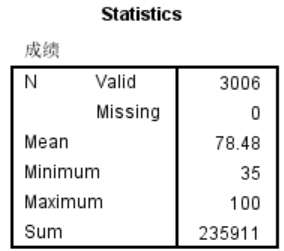


图4统计分析结果

**（二）学生成绩特征描述**

在“Frequencies:Statistics”对话框中，勾选“Quartiles”，设置“Cut points for5 equal groups”，“Percentile(s)”也设为5，可以得到学生成绩的Percentiles。同理，可以描述和度量学生成绩的集中趋势、离散程度和分布形态，在以上对话框中勾选所有复选框，SPSS Statistics便可自动计算出各自结果。

结果显示，平均值，可以认为此次考试试题难度适中，比较符合学生的实际水平，也表明教师在教学过程中能够照顾到多数学生对知识的理解能力。标准差，表明此次考试的成绩离散程度稍微偏大。从分数的最大值（100分）和最小值（35分）也可以看出，学生的成绩范围较大（65分），导致标准偏差偏大。因此，教师在平时的教学中应多关注成绩较低的学生，并对其针对性的指导，避免这些学生的成绩与班级平均分出现太大差距。

通过进一步分析探索分析，可以比较此次考试男、女生成绩的差异，借助SPSS Statistics，依次选择“Analyze—Descriptive Statistics—Explore”，将“成绩”设置为Dependent List，将“性别”设置为Factor List。“Plots”按钮下级对话框的Descriptive只选择“Histogram，其他默认设置，即可得到描述性结果。

结果显示，，即两者大体相等，说明离散程度相当，但女生的平均分相对较高，表明女生对知识的理解和掌握能力较好。从男女生成绩的直方图中，也可以直观地看出女生的成绩比男生要高，如图5所示。作为任课教师，也可根据学生的性别进行差异化指导和教学，确保男女学生成绩均衡。

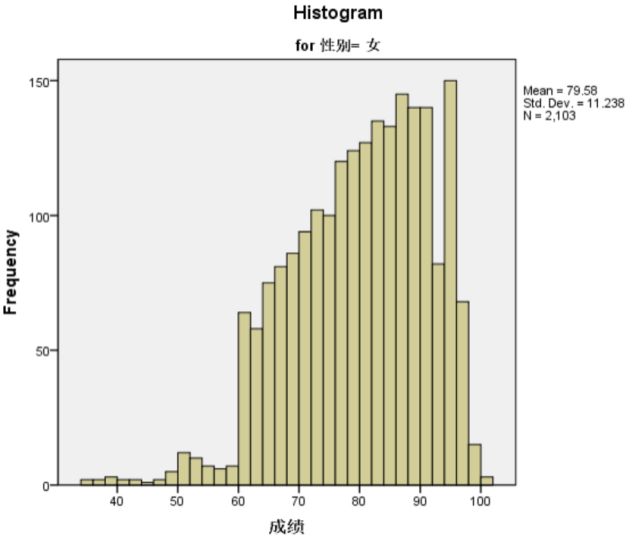
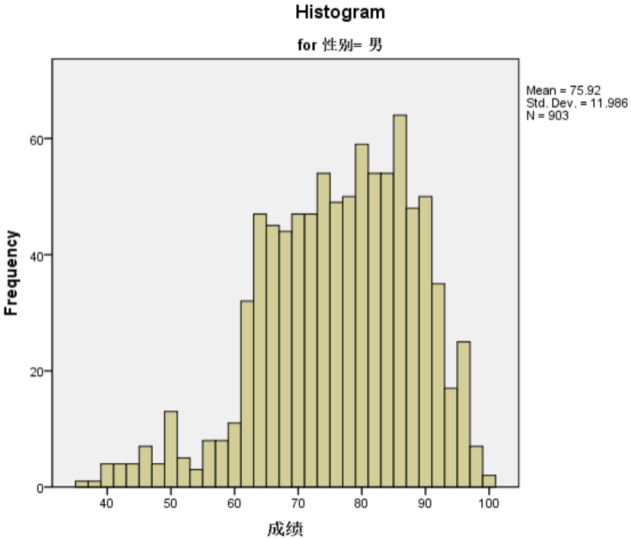


图5男、女生成绩直方图

此外，还可分析不同系别或不同专业学生成绩的差异。如果学生成绩大体相等，则越小越好，说明学生成绩比较集中，与的差异也比较小；如果和均有明显的差异，则可根据标准差系数的大小来观察其相对离散程度，标准差系数越小越好，表明大部分学生能同步发展，也表明教师的教学方法能够适合多数学生。

**（三）频率分布**

根据学生的考试成绩，借助SPSS Statistics，可用带正态曲线的直方图来表示学生成绩的总体分布。还可用Q-Q图检验学生的考试成绩，如图6所示，可以看出下端虽有较多点偏离直线，但粗略判断也接近正态分布。

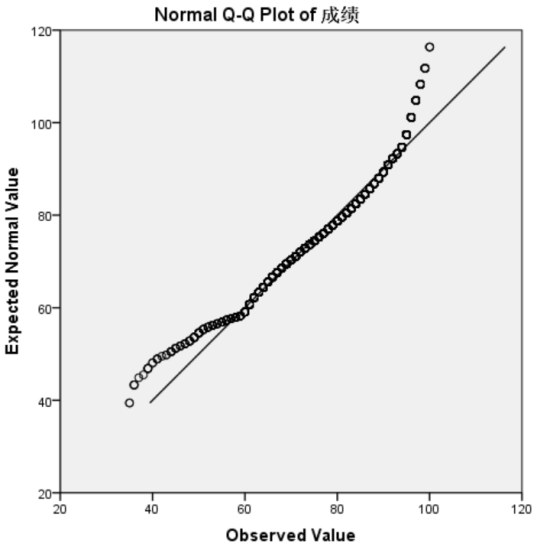


图6 学生成绩Normal Q-Q Plot

根据前面结果可知，本次考试学生成绩的*bs*=-0.590，*bk*=0.157，可判断成绩曲线是呈负偏态分布且呈正态分布，这与带正态曲线的学生成绩直方图是一致的，也与图6中“Normal Q-Q Plot of 成绩”检验结果粗略判断其接近正态分布是一致的。由于*bs*<0，表明高分人数较多，低分人数较少；成绩接近正态分布，表明试题难度比例比较合理，能够与学生的学习能力相适应，也表明教师在命题时能够较好地把握学生的学习水平。因此，教师在命题时应该正确把握试题的难度，控制试题难度比例，一则保证学生不会因为试题难度过大，导致平均分太低，二则又能考察拔尖学生的能力，体现出差距，这样，学生的成绩也就更趋向正态分布。

**（四）有效性分析**

通过以上分析可知，使用特征描述的方法分析学生成绩，以此评价教师命题的科学性，除了使用传统简单的评价指标外，还能借助软件从描述和度量学生成绩的百分位值、集中趋势、离散程度和分布形态来进一步评价，使用这种方法能避免人工计算可能带来的错误，又能有效提高成绩的分析效率。同时，简易的操作步骤、可视化的交互界面、清晰的统计分析结果以及结果的可验证性，使其深受多数用户的欢迎，因此，也成为初级用户选择数据挖掘方法的首选。

综上所述，根据教师命题科学性的各项评价指标，使用数据挖掘的特征描述方法，分析学生成绩的总体水平特征分布与频率分布，并评价教师命题的科学性。整个过程，借助SPSS Statistics软件对学生成绩进行度量与描述，实现自动分析、展示图形以及验证结果，分析结果能够辅助教师科学命题，提高试题的编制质量与教学质量。

**参考文献：**

[1]韩家炜.数据挖掘：概念与技术[M].北京：机械工业出版社,2007:10.

[2]丁伟，李政，于昕.数据挖掘在政府采购中的应用研究[J].中国政府采购,2014，(07):72-73.

[3]宋赟.基于数据挖掘技术的学生成绩分析[D].重庆：重庆师范大学,2009：8.

[4]尹群.2015年吉安市环境空气质量现状评价与分析[J].江西化工,2016，(04):95-98.

[5]方洪鹰.数据挖掘中数据预处理的方法研究[D].重庆：西南大学,2009：13.

**An Evaluation of the Scientificity of Test Papers Based on Students' Score Analysis**

**——Application of Data Mining and Feature Description**

CHEN Xihua1, HUANG Haining1, HUANG Peijie2

1. Guangzhou College of Technology and Business, Guangzhou, Guangdong 510850, China;   
   2. South China Agricultural University, Guangzhou, Guangdong 610642, China）

**Abstract:** Under the method of Excel functions and the data analysis method of feature description, and with the SPSS statistics software, measures of dispersion, discrimination, difficulty, the distribution of the characteristics of the overall level and the frequency distribution of the students’ scores are analyzed in order to evaluate the scientificity of the test papers and help teachers understand the distribution of students’ scores and their comprehension in courses. The result can also assist teachers to make tests and improve the quality of test questions as well as teaching.

**Key words:** data mining; characteristic description; SPSS statistics; test paper quality; scientific analysis

（责任编辑：孙巧云）

1. 收稿日期：2018年05月25日

   作者简介：陈喜华（1985-），男，广东潮州人，助理研究员，硕士学位，研究方向：教育教学管理，计算机应用。

   黄海宁（1991-），女，广东广州人，硕士学位，研究方向：学科教学。

   黄沛杰（1980-），男，广东潮州人，副教授，博士，研究方向：人工智能、自然语言处理。

   基金项目：2017年广州工商学院质量工程--高等教育教学改革项目（ZL20171138）研究成果；广东省教育研究院民办教育研究基地（广州工商学院）2017年科研课题《民办高校学科、专业、课程、教材建设与教学改革研究》（编号：GMJ201707）。 [↑](#footnote-ref-0)