

# 用相关性原则甄别科技论文中的数据错误

张凤周瑜李建英

《大连理工大学学报》编辑部, 116024, 辽宁大连

**摘要** 数据图表等是科技论文的基石,但因作者经验、时间、不严谨等原因而导致错误时有发生,对论文学术质量造成了影响。编辑运用科技编辑方法论中的相关性原则,凭借丰富的经验和直觉可以发现并改正此类错误。

**关键词** 科技论文;图表;相关性原则;勘误

**Discrimination of data errors in sci-tech papers by relativity principle** //ZHANG Feng, ZHOU Yu, LI Jianying

**Abstract** Data, figures and tables are the key contents of sci-tech papers. Because of the deficiency of writing experience, time and precise attitude, the data errors occur frequently, which reduces the quality of the papers. Combining with the plenty of experience and professional intuition, the authors hold that, the editors of sci-tech periodicals can use the relativity principle in editorial methodology to reduce the errors.

**Key words** sci-tech paper; figure and table; relativity principle; discrimination of error

**Authors' address** Editorial Office of Journal of Dalian University of Technology

期刊质量包括学术质量和出版质量2个方面。学术质量通常用一定的客观指标如影响因子、被引频次、检索系统收录情况、基金论文比、平均引文数等衡量<sup>[1]</sup>,但这些均为外在的评价指标,难以定量评价论文的学术水平。作为科研成果的主要载体,科技论文具有公式图表多、符号单位多、数据多、学科交叉性强、术语多而新、分析讨论逻辑性强等特点,在规范表达的基础上,数据的准确性尤为重要。而许多科技论文出自博士生之手,他们或受限于经验、时间,或者学术作风不够严谨,导致在讨论部分数据错误时有发生。编辑在工作实践中发现,运用系统方法论中的相关性原则对甄别科技论文中的数据与图表错误非常有效。笔者结合具体实例,对此进行探讨。

相关分析研究现象之间是否存在某种依存关系,并对具有依存关系的现象探讨其相关方向以及相关程度,而相关性原则是相关分析思想的总结。相关性分析通常应用于统计分析,尤其是数据分析<sup>[2]</sup>,而相关性原则则广泛应用于经济、医学、计算机等领域,用来预测某一时间段的进展或者结果<sup>[3-5]</sup>。

本文利用编辑方法论中相关性原则中的一一对应、逻辑关系等原理<sup>[6]</sup>对科技论文中的错误进行剖析。

## 1 从组内数据相关关系判断错误

任何事物都是有内在联系的,实验数据间更是有着密不可分的关系,尤其是在同一组数据中。组内数据常见的错误有以下几种:

- 1) 数据重合、交叉定义;
- 2) 相同属性的数据数量级差别很大;
- 3) 相同属性的数据小数位数或有效数字不一致;
- 4) 相同属性的数据结果计算规则不同;
- 5) 计算错误或者笔误。

**例1** 河流功能分级中,按 pH 大小打分如下: 6.5~7.5(4分), 6.0~6.5(3分), 5.5~6.0(2分), 4.5~5.5(1分), 4.5以下(0分)。

**分析** 从此例中可见同一 pH 得分不同,说明此种分级方法混乱。其原因是未考虑数据间的交叉关系。正确的应为: 6.5~<7.5(4分), 6.0~<6.5(3分), 5.5~<6.0(2分), 4.5~<5.5(1分), <4.5(0分)。注意定义分段函数时,不要重复定义某点的函数值。

**例2** 判别值与真实值之间的关系见表1。

表1 计算结果

真实值	判别值	判别值与真实值差值
0	0	0
5	0	5
10	17	7
15	25	10

**分析** 由此表可以看出,第2行的数据差值“5”有误,正确的应为“-5”。说明计算数据间的差值时没有考虑到正负号或者减数和被减数的关系。注意要与绝对误差和相对误差相区分,绝对误差是测量值与真实值之差的绝对值,相对误差为绝对误差所占真实值的百分比,一般在计算时大多采用的是绝对误差和相对误差,计算单纯的前后数字差值时,不能用绝对值,要考虑正负号。另外,如化学元素分析中,各组精确计量的,相加后结果必须为100%。

**例3** 实验研究轴压力系数、体积配箍率对钢筋高强混凝土延性的影响,具体结果见表2。

**分析** 从试件编号和试验参数的对应关系可以发现,试件编号6是错误的,因为试件编号的第1组数对应剪跨比,正确的编号应为 SRC 20-16-115(试件1~6

的剪跨比均为 2.0);此外,体积配箍率的第 5 项是错误的,因为根据试件编号判断,其相应值应为 2.2。编辑数据表时一定要注意其间的相关关系。

表 2 钢筋高强混凝土短柱的位移延性系数

序号	试件编号	立方体强度/MPa	体积配箍率	轴压力系数	位移延性系数
1	SRC 20-22-125	113.2	2.2	0.73	2.00
2	SRC 20-16-125	93.7	1.6	0.73	1.80
3	SRC 20-12-125	110.9	1.2	0.73	1.65
4	SRC 20-08-125	100.0	0.8	0.73	2.06
5	SRC 20-22-115	118.7	1.6	0.67	2.09
6	SRC 15-16-115	85.6	1.6	0.67	1.63

## 2 前后数据间的相关关系

数据间的逻辑关系构成了数据的联系,一篇论文中前后组实验数据往往是有内在联系的。数据关系相同,则得到的相同属性的数据应该有相同的特征,特征外的值都应该是错误的。组间数据常见错误有以下几种:1)实验参数与实验结果中采用的实验参数不对应;2)前组实验结果与引用前组数据的后组实验结果参数不对应;3)图表规律、图文规律不相符。

例 4 图 1 为平均孔径对孔径分布的影响,图 2 是根据图 1 中不同的孔径分布函数模拟得出的通量变化。可见在膜的其他几何参数都相同的情况下,随着膜平均孔径的增大,膜渗透通量也显著增大。

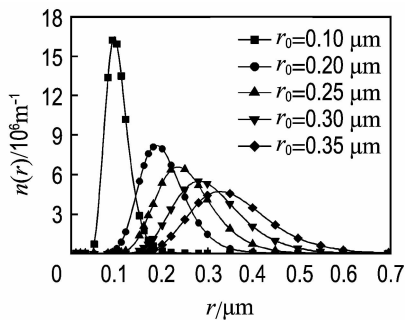


图 1 平均孔径对孔径分布的影响

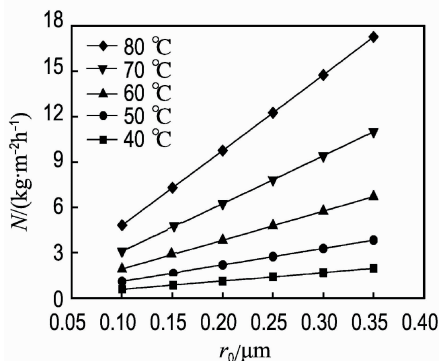


图 2 平均孔径对通量的影响

分析 有些数据间是有因果关系的,此例中图 2 是根据图 1 中部分实验数据进一步实验得出的结果,采用的平均孔径分别为 0.10、0.20、0.25、0.30、0.35 μm,而文中的图 2 里多了 0.15 μm 这一组数据,显然是不合理的。

例 5 如图 3 可见,业务代表 E5 的重要性最高。

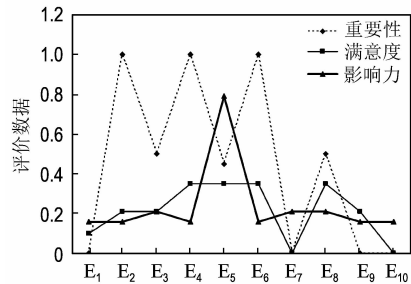


图 3 对 10 位业务代表个人特征评价结果

分析 由此图可以看出,业务代表 E5 的影响力最高,而重要性最低,与正文叙述不符,经与作者核实发现是图的图例反了。

另外,要注意数值的正负情况,即坐标轴标值与图例的符号要综合考虑。

由上述实例可见,一组数据内,几组数据间,图表间,图图间及图文间都是有逻辑关系的;因此,严密的逻辑思维能力是编辑不可缺少的,要用这种能力发现错误,改正错误。

## 3 结束语

编辑工作中运用相关性原则,凭着严谨的工作作风,通过经验的积累,可以帮助作者极大地减少数据错误,提升论文的学术质量。

## 4 参考文献

- [1] 陈冠初. 我国科技期刊的评价问题[J]. 编辑学报, 2004, 16(4): 240-242
- [2] 郑明, 陈子毅, 汪嘉冈. 数理统计讲义[M]. 上海: 复旦大学出版社, 2007
- [3] 王胜本, 徐晓肆, 陈瑛. 中国股票市场相关性分析[J]. 山东科技大学学报: 自然科学版, 2008(6): 96-101
- [4] 刘玉梅, 张德芹, 李玉红, 等. 2 型糖尿病与抑郁症生物学相关性研究进展[J]. 中国康复, 2008, 23(3): 198-199
- [5] 范植华. 数据相关性理论及其在并行处理中的应用[M]. 北京: 科学出版社, 1999
- [6] 钱文霖. 科技编辑方法论研究[M]. 武汉: 华中理工大学出版社, 1998

(2010-07-21 收稿; 2010-09-26 修回)