

科技期刊刊文相关分析中存在的问题

许育彬

西北农林科技大学《麦类作物学报》编辑部,712100,陕西杨凌

摘要 科技期刊刊文中相关分析方面存在诸多问题,如稿件中相关分析的应用远离研究主题,完全正相关系数的取值不正确,相关分析表信息重复,没有标注样本数量。科技期刊编辑应重视对这些问题的审查,并尽最大可能在稿件审阅阶段予以解决,同时可通过补白等方式将一些统计方面知识的推介给广大读作者,促进稿件质量提高。

关键词 科技期刊论文;相关分析;问题

Some problems of correlation analysis in papers of sci-tech journals//XU Yubin

Abstract There are some problems of correlation analysis in many papers of sci-tech journals, such as its use irrelevant to research themes, wrong perfect correlation coefficient value, information repetition in the correlation coefficient table, and absence of simple size in the note under the correlation coefficient table. The problems should be paid attention to by editors, and be solved in the paper review while some knowledge of statistics may be introduced to readers and authors by proper forms such as filler.

Keywords papers of sci-tech journal; correlation analysis; problem

Author's address Editorial Department of Journal of Triticeae Crops, Northwest A&F University, 712100, Yangling, Shanxi, China

DOI:10.16811/j.cnki.1001-4314.2016.04.008

相关分析是研究现象之间的依存关系及其相关方向和程度的一种统计方法,经常在科研中被用于探究随机变量之间的相关关系^[1]。有关科技期刊论文中统计分析存在的问题已有很多报道^[2-7],但涉及相关分析的研究很少,而且基本上属于概念问题,如常用相关系数、复相关系数和复确定系数的区分等^[7]。在多年的《麦类作物学报》编辑实践中,笔者发现,很多稿件中相关分析应用存在诸多问题;在中国知网以“相关分析”为主题对科技期刊2015年刊文进行随机抽查,发现此类问题具有普遍性;因此,很有必要对这些问题进行归纳分析,以供编辑同人参考。

1 相关性分析存在的问题及分析

1.1 相关分析的必要性 科技论文的结论是建立在坚实可信的数据及其合理分析之上的^[8],只有合理、有效的数据分析才能为结论的得出提供更多有用的论据,进而解答研究中提出的问题或假设^[9],也就是说数据分析是为研究主题服务的。在一些文章中,相关

分析没有完全考虑试验设计中因素的作用,而是单纯地对所考察的指标或性状间的关系进行相关分析,这种相关性并不能反映试验因素对所测定指标或性状的影响。

例如《核农学报》2015年第6期《NaCl处理对海滨锦葵N、P和Na⁺、K⁺含量及其化学计量特征的影响》和《山西农业科学》2015年第5期《种植密度对不同玉米品种农艺、光合性状及产量的影响》的相关分析均是抛开试验因素(NaCl、种植密度)来分析所测指标或性状间的相关性,远离了研究主题;因此,编辑应对稿件中相关分析的必要性进行审阅和合理修改,以去芜存菁。

1.2 相关系数最大值 相关系数在[-1,1]区间取值,其最大绝对值应为1,表示变量间完全相关;但在一些期刊的刊文中,相关分析表为了保持一致的小数位,完全正相关系数写成1.0、1.00、1.000等整数位为1的小数,如《植物营养与肥料学报》2015年第1期第99页的表1、《生态学报》2015年第12期第3985页的表2、《环境科学学报》2015年第5期第1370页的表2等。

在数学上,1与1.0、1.00、1.000等小数的大小相同,但所表示的意义不同^[10],1.0、1.00、1.000等有可能是由大于或小于1的值得来的,如0.995~1.004都可以进位成1.00,如果完全正相关用1.0、1.00、1.000等小数表示,就很容易使读者产生误解,因而完全正相关系数应是1。

1.3 多元相关分析表的简化 多元相关分析表通常包含2个以上变量,它们之间的相关系数在表中的排列呈三角形,三角形长边即对角线上的数据为1(表1)。由于一些统计软件的数据输出结果在表中的排列为沿对角线呈互为镜像的2个三角形,数据重复,因此在数据整理时只需要列出一个三角形的数据。

调查发现,部分期刊的论文依然将2个三角形数据列出,如《农业机械学报》2015年第1期309页的表2。此外,在变量数量较多的情况下,多元相关分析表往往很大,会给排版造成一定困难,需要对表格进一步简化,删除多余的信息。对表1的观察发现,在保留下三角数据后,第2行和最右列的数据均剩下对角线上的1,因此完全可以取消这一行、一列,缩小表格,为表格合理排版创造条件。

表1 多元相关分析表的原始格式

	X_1	X_2	...	X_n
X_1	1			
X_2	V_{12}	1		
\vdots	\vdots	\vdots		
X_n	V_{1n}	V_{2n}	...	1

注: X 和 V 分别表示变量及其相关系数。

1.4 相关系数的显著性注解 科技期刊论文相关分析表中,相关系数显著性多用单、双星号标注,以分别表示达到0.05和0.01显著水平,再在表后用“* : $P < 0.05$; * * : $P < 0.01$ ”“*和* *分别表示显著($P < 0.05$)和极显著($P < 0.01$)”等之类文字加以注解,也有不注解的(如《河北农大学学报》2015年第1期有5个相关分析表,其中3个未注解);也有不用星号标注,只在表后附注相关系数达到0.05或0.01显著水平的临界值。

然而,很少见到在相关分析表后对样本数量进行说明的。不同样本量下相关系数的显著性临界值不同,样本量越大,临界值越小^[1],对相同的相关系数以不同的临界值来衡量,得到结论不同;因而为了保证表格信息的完整性和自明性,很有必要对样本数量在相关分析表后加以说明,尤其是在同一篇文章多个相关分析表或同一表格中不同相关分析的样本量不一致的情况下就更有必要,否则会出现信息混乱,易导致读者误解数据结果。

例如《麦类作物学报》2015年第8期第1104页表3中半冬性小麦和春性小麦的样本量分别是6和9,如果仅用“* : $P < 0.05$; * * : $P < 0.01$ ”注解,读者看到后会发现有些较大的相关系数不显著,有些较小的相关系数反而显著,有可能认为显著性标注有误;因此,增加样本量标注是对“* : $P < 0.05$; * * : $P < 0.01$ ”注解的有力补充。

2 应对策略

稿件中出现相关分析问题的原因与作者、审稿专家和编辑自身数理统计知识不足或疏忽大意有很大关系,因此有必要从这三者的角度去解决此类问题。

首先,编辑要不断学习、积累和掌握相关数理统计知识,在审稿过程中依靠这些知识,认真处理稿件,尽量将这些问题在稿件退修时予以解决,如果有不懂的地方也可以咨询专门从事统计分析研究的专家。这样,不仅可增加相关知识,也有助于保证稿件质量。

其次,在稿件送外审时,应要求审稿专家在对稿件学术水平进行审阅和评价的同时,对其数据统计分析也进行评审。这在审稿单设计时就应作为一个重要项

目加进去^[11]。

其三,在《投稿须知》中应强调作者关注稿件中的数理统计问题,也可以利用期刊补白、期刊网站、期刊微信公众号、QQ群等将一些统计学知识推介给读者和作者,对投稿中一些共性较强的数理统计问题可像《植物生态学报》那样进行专门探讨^[9]。

在这3个方面中,编辑的责任重大,编辑责任心强、要求高了,稿件审阅时发现问题的概率也就高,送审稿件时对审稿人的选择也就严格了,有助于在稿件评审阶段解决更多问题。

3 结束语

以上相关分析问题在科技期刊论文中具有普遍性。由于研究对象数量很大,笔者采取随机抽查方式进行数据采集,若在抽到的期刊某一期中发现问题,就不再对当年的其他期进行调查;但在随机抽取的300多种期刊中发现问题的期刊比例超过20%,在一些期刊同一期中同样的问题甚至多处出现,因此对刊文中相关分析的审核应引起编辑同人高度重视。

文章写作过程中,《西北农业学报》编辑部郭柏寿老师给予很多建议和帮助,在此表示诚挚谢意。

4 参考文献

- [1] 袁志发,周靖芊. 试验设计与分析[M]. 北京:高等教育出版社,2000
- [2] 张巨稼. 略论科技期刊论文中统计方法的误用[J]. 中国科技期刊研究,2012,23(3):413
- [3] 夏爱红,景桂英,是雅蓓. 农业期刊论文中试验设计及统计方法存在的问题及分析[J]. 编辑学报,2002,14(1):23
- [4] 邵晓军. 科技期刊编辑应注意统计描述中的错误[J]. 赤峰学院学报(自然科学版),2011,27(8):208
- [5] 王音. 农业和生物学期刊应重视对表格中数据统计分析结果的审校[J]. 编辑学报,2010,22(5):413
- [6] 吴学军,廖粤新. 科技期刊编辑应注重统计数据的审核[J]. 编辑学报,2010,22(5):416
- [7] 郝拉娣,关伟,于化东. 科技期刊中生物统计方法的正确表达[J]. 编辑学报,2004,16(2):109
- [8] 匡少平,王亨军. 科技期刊编辑过程中的数据分析问题[J]. 编辑学报,1994,6(2):71
- [9] 胡凤琴,李硕,牟溥. 如何写好数据分析和几个相关的统计问题[J]. 植物生态学报,2013,37(6):583
- [10] 戎松斛. 如何正确理解小数的意义[J]. 小学教育,2013(13):20
- [11] 王庆霞,韩蕾,刘浩芳,等. 科技论文外审审稿单设计要点及实例[J]. 编辑学报,2013,25(S1):15

(2015-12-31 收稿;2016-04-18 修回)