

医学研究中定性资料假设检验方法误用分析

李 明 贾秀琴 陈嫚茵 李晓萍 李耀芳 吴正治
(深圳市中西医结合临床研究所, 广东 深圳 518031)

[摘 要] 通过解析医学研究论文中使用定性资料假设检验方法的常见错误, 介绍了定性资料常用假设检验方法, 指出在使用统计方法时应掌握其适用条件, 避免误用。
[关键词] 定性资料; 假设检验; 误用
[中图分类号] G304 [文献标识码] A

Analysis of Common Misuses of Significance Test of Qualitative Data in Medical Papers

LI Ming, JIA Xiu-qin, CHEN Man-yin, LI Xiao-ping, LI Yao-fang, WU Zheng-zhi
(Shenzhen Clinical Institute of Integrated Medicine, Guangdong Shenzhen 518031)

[Abstract] Common misuses of significance test of qualitative data in medical research papers were analyzed, meanwhile, common significance test methods of qualitative data were introduced, and the authors should master the applying conditions of statistic methods were pointed out.
[Key Words] Qualitative Data; Significance Test; Misuse

近年来医学统计学在医学研究中的应用日益受到广大医学工作者的重视, 但是在医学统计方法的实际运用中也常常出现一些问题, 例如忽视方法的应用条件、或是选择的方法不恰当等, 从而影响研究结论的正确性和可信性。为减少这一现象, 笔者对医学研究论文中定性资料假设检验方面常见的错误分析如下。

1 医学论文中常见的定性资料假设检验错误

1.1 四格表资料

例 1: 某临床研究比较尿激酶股动脉注射与手术取栓治疗 ADVT 疗效, 结果见表 1。该文作者根据 $\chi^2=4.288$, $P<0.05$, 认为两组治愈率差异有显著性。

表1 尿激酶股动脉注射组与手术取栓组比较

组别	n	临床治愈人数(η/%)
股动脉注射组	15	3(20.00)
手术取栓组	16	9(56.25)

错误分析: 作者忽视了本资料样本含量 $n<40$, 不符合四格表资料 χ^2 检验适用条件, 应选择四格表资料的 Fisher 确切概率计算法, 经计算得出本资料 $P=0.066$ 。故尚不能认为两组治愈率差异有显著性。

例 2: 某研究测定大肠癌组织与正常大肠组织的 VEGF 表达, 所得资料见表 2。结论: 大肠癌组织中 VEGF 表达显著高于正常大肠组织。

[收稿日期] 2007-01-27
[作者简介] 李明(1970-), 女, 湖北省荆州市人, 讲师, 主要从事中西医结合基础研究工作。

表2 两组VEGF表达比较

组别	n	阳性	阴性
癌组织	46	39	7
正常组织	12	0	12

$\chi^2=31.057, P < 0.01。$

错误分析：本资料适用四格表资料 χ^2 检验，但作者未注意有一格子（正常组织阴性表达栏）理论频数 $T=3.93$ ，而当 $1 \leq T < 5$ 且 $n \geq 40$ 时，需计算校正 χ^2 值。经计算，本资料校正 $\chi^2=27.328, P=0.000$ 。虽然结论与原文相同，但原文统计方法选择有误。

例 3：某作者研究原发性肝癌（PHC）的实验室诊断方法，对经过病理诊断证实的 PHC 患者同一份血清标本分别检测 AFP 和 TSGF，比较两种诊断方法的结果有无差别，结果见表 3。经 χ^2 检验（未给出检验统计量）， $P>0.05$ ，结论为 2 种检验方法的差别无统计学意义。

表3 PHC患者AFP和TSGF检测结果比较

AFP检测	n	TSGF检测	
		+	-
+	84	76	8
-	28	22	6

错误分析：该研究采用配对实验设计，因此假设检验方法应该用配对设计四格表资料的 χ^2 检验，并且本例 $b+c=30<40$ ，故计算校正的 $\chi^2=5.633, P=0.018$ ，结论为 2 种检验方法的差别有统计学意义。该结论与原文的结论正好相反。

1.2 2×k 表资料

例 4：有人研究脾虚泄泻与食物过敏的关系，探讨饮食调整改善脾虚泄泻症状的价值，所得结果见表 4。经 χ^2 检验， $\chi^2=57.263, P<0.01$ ，认为饮食调整组改善作用优于对照组。

表4 两组患者疗效比较

组别	n	完全缓解	明显缓解	轻度缓解	无效
饮食调整组	57	13	37	3	4
对照组	57	1	9	13	34

错误分析：表 4 属单向有序 $2 \times k$ 表资料（疗效按照从好到差顺序排列），拟比较两组间的疗效差异，不适合选择 χ^2 检验进行统计分析，因为 χ^2 检验与“疗效分级”的有序性之间没有任何联系，如果我们对表 4 中的任何两列数据进行对换， χ^2 值也不会改变。应该选择与有序性有联系的秩和检验或 Ridit 分析。经 Wilcoxon 秩和检验， $u=7.225, P=0.000$ ，差异有显著性。经 Ridit 分析， $u=6.859, P=0.000$ ，差异有显著性。虽然经过正确的统计方法分析的结论与误用 χ^2 检验结论一致，但还是应选用正确的统计方法。

1.3 R×C 表资料

例 5：某临床研究观察生肌散治疗褥疮的疗效，经过两个疗程的治疗，所得结果见表 5。经 χ^2 检验， $\chi^2=48.239, P<0.01$ ，认为治疗越早，疗效越好。

表5 生肌散治疗褥疮的疗效观察

褥疮分期	治疗结果			合计
	治愈	好转	无效	
I	57	12	0	69
II	12	9	6	27
III	4	8	12	24

错误分析：该资料的 2 个变量“褥疮分期”和“治疗结果”都是有序的（双向有序），即有等级关系，而且属性不同。分析这类资料不适合选择 χ^2 检验进行统计分析，因为 χ^2 检验与 2 个变量的“有序性”之间没有联系，应选用与 2 个变量的“有序性”有联系的等级相关或线性趋势分析等假设检验方法。本资料经 Spearman 等级相关分析， $r_s=0.604$ ，

$P < 0.01$, 可认为应用生肌散治疗褥疮越早疗效越好。虽然结论与原文相同, 但原文统计方法选用有误, 并且即使选用 χ^2 检验, 因为本文数据有 1/5 以上的格子数理论频数 $T < 5$, 也不宜采用直接计算得到的 $\chi^2 = 48.239$, 可选用合理合并行、合并列、删除行、删除列等方法调整数据后再计算。

2 讨论

定性资料是观测每个观察单位某项指标质的状况所得的资料。例如某市全体员工按职业(工人、农民、军人……)来记录每个人的情况所得的资料。在定性资料中要注意区分“有序资料”, 这是一类指标质的不同状况之间在本质上有数量大小或有先后顺序之分的定性资料。例如矽肺患者按肺门密度级别(+、++、+++)来划分所得的资料。在统计分析过程中, 要根据资料的特点和假设检验的适用条件选用统计方法。

2.1 四格表资料假设检验方法的选择

对两个样本率或构成比进行比较时, 常选用 χ^2 检验, 包括基本公式、四格表资料 χ^2 检验的专用公式及四格表资料 χ^2 校正公式计算法^[1]。

2.1.1 当 $n \geq 40$, 且所有格子的理论频数 $T \geq 5$ 时, 直接用专用公式或基本公式计算。

2.1.2 当 $n \geq 40$, 但出现 $1 \leq T < 5$ 时, 需计算校正 χ^2 值。

2.1.3 当 $n < 40$, 或有任一格子 $T < 1$ 时, 需要用 Fisher 确切概率法。

对于配对设计比较两种处理的结果有无差别时, 采用配对设计两样本率比较的 χ^2 检验, 注意当 $b+c < 40$ 时, 须计算校正的 χ^2 值。

2.2 $2 \times k$ 表资料假设检验方法的选择

对于列为无序资料的 $2 \times k$ 表, 可用 χ^2 检验法。对于列为有序资料的单向有序 $2 \times k$ 表, 应该选择与有序性有联系的秩和检验或 Ridit 分析^[2]。

2.3 $R \times C$ 表资料假设检验方法的选择

当 $R \times C$ 表的行与列皆无序时, 可用 χ^2 检验法。但要注意以下事项。

2.3.1 $R \times C$ 表的 χ^2 检验要求理论频数不宜太小, 一般不宜有理论频数小于等于 1, 且 $1 < T \leq 5$ 的格子数不宜超过总格子数的 1/5。

2.3.2 如果以上条件不能满足时, 不能直接作 χ^2 检验, 此时可以采用以下方法: ①增加样本含量, 当样本含量增加时必然会引起理论频数的变化, 此时可能会使得数据满足 χ^2 检验的条件。②删去某行或某列, 删去理论频数较小的行或列使得资料满足 χ^2 检验的条件, 但删去某行或某列往往会减少收集到的数据信息。③合理地合并部分行或列。④用 Fisher 确切概率法计算^[3]。

当 $R \times C$ 表的行与列之一有序时(即单向有序 $R \times C$ 表), 应该选用秩和检验或 Ridit 分析等方法; 对于双向有序且属性不同的 $R \times C$ 表资料, 若希望弄清两有序变量之间是否存在线性相关关系或存在线性变化趋势, 应选用定性资料的相关分析(如 Spearman 等级相关分析)或线性趋势检验; 对于双向有序且属性相同的 $R \times C$ 表资料, 为考察两种方法检测的一致性, 应选用 Kappa 检验^[4]。

综上所述, 在分析医学研究数据中的定性资料时, 切忌只关心 P 值的大小而忽视统计分析方法的适用条件, 盲目套用公式而造成统计分析方法的误用。在实际应用中, 必须根据资料所具备的条件及分析目的, 合理选用统计分析方法。

(参考文献)

- (1) 刘桂芬. 卫生统计学(M). 北京:中国协和医科大学出版社, 2003. 84-87
- (2) 何清波, 苏炳华, 钱亢. 医学统计学及其软件包(M). 上海:上海科学技术文献出版社, 2002. 186-187
- (3) 于浩. 医学统计学(M). 北京:中国统计出版社, 2005. 91
- (4) 徐勇勇, 倪宗瓿. 医学统计学与卫生统计学(M). 北京:科学技术文献出版社, 2002. 104