

医学科研论文初审时统计学错误的发现技巧

邓强庭

(第三军医大学 学报编辑部,重庆 400038)

摘要:医学统计学作为医学科研的一个重要工具,在医学科研论文中有着广泛的应用。处理医学科研论文在初审时迅速发现论文中统计学方法的不妥或错误,做出直接退修或者退稿的处理,可以减少不必要的后续审稿流程,提高稿件处理速度,降低办公成本。文中总结医学统计学错误主要存在样本量不达标、统计学方法的选取不恰当、 χ^2 检验的基本条件不满足、检验水准确定不正确等几个方面,而这些错误是在初审时比较容易发现的,文中就初审环节对这些统计学方法的审核交流一些方法和技巧。

关键词:医学科研论文;统计学;假设检验

中图分类号:G237.5

文献标识码:A

文章编号:1672-5379(2010)05-0268-02

医学统计学是医学科研的一个重要工具,在医学科研论文中有着广泛的应用。在处理医学科研论文时,对统计学方法的审核是非常重要的环节,一篇医学科研论文中统计学方法的正确与否直接影响着论文的科学性和结论的正确性。医学论文编辑们每天处理大量的论文,如果能在初审时迅速发现论文中统计学方法的不妥或错误,从而做出直接退修或者退稿的处理,则可以减少不必要的后续审稿流程,提高稿件处理速度,降低办公成本。现就初审环节对统计学方法的审核交流一下经验。

一、样本量是否达到要求

样本量的本质其实是反映科研设计中“可重复”的基本原则,对科研成果的科学性有至关重要的影响。严格意义上讲,样本量应该根据研究内容的特征进行查表或者计算得到^[1],如,常用的两样本均数检验的样本量计算公式为:

$$N = \left[\frac{(Z_{\alpha/2} + Z_{\beta})\sigma}{\delta} \right]^2 (Q_1^{-1} + Q_2^{-1})$$

式中, $Z_{\alpha/2}$ 为检验水准 α (即犯第 I 类错误的概率)对应的双侧的 Z 值,而 Z_{β} 为犯第 II 类错误的概率 β 对应的单侧的 Z 值, Q_1 、 Q_2 为两组样本的样本比例, δ 、 σ 分别为容许误差和总体的标准差,而事实上总体的标准差是几乎不可能得到的,若要计算只能用预试验来估计。这样的计算方法相当繁琐,而且其中很多参数在实际的科研中往往难以得到确切值,很多科研工作很难完全按照公式计算样本量,但是也应该有一个经验性的大概范围,例如用 χ^2 检验时,

总例数 n 应该不少于 40,因为在 $n < 40$ 时, χ^2 的所有公式均不可用了,只能选取其他的检验方法,如确切概率值法^[2]。在审稿时若发现论文使用了 χ^2 检验,但 n 却少于 40,就可以认为该论文的统计学方法存在错误,应该退给作者,建议重新选取统计学方法重新计算或者扩大样本量重新计算。例如某来稿中,作者写明使用的统计学方法是 χ^2 检验,但从表中数据可知总例数只有 30,不符合要求,见图 1。

1.4 统计学方法采用 χ^2 检验。

2-结果。

2.1 两组患者治疗效对比(表 1)。

表 1 两组患者治疗效果。

	痊愈	显效	好转	无效
治疗组	8	5	1	4
观察组	4	4	3	4

图 1 某来稿中样本量未达到 χ^2 检验的要求

二、统计学方法的选取是否恰当

把计量资料用 χ^2 检验或者把计数资料用 t 检验这类的错误是很明显的,通常很容易发现,但是除此之外还应该注意,往往有研究者凡是遇到计数资料都一概用 χ^2 检验,这是不合适的,计数资料中,如果研究指标分类是无序的,通常应用 χ^2 检验没有问题,例如观察疗效分为“治愈”和“未治愈”,或者研究避孕方式将其分为“口服避孕药、节育手术、使用

收稿日期:2009-11-06

作者简介:邓强庭(1984—),男,四川中江人,第三军医大学助理编辑,研究方向:医学论文的统计学、科技期刊数字化。

避孕套”等。但是如果研究指标并非互相独立而是有递进关系,例如将疗效分为“无效、好转、显效、治愈”或者临床检验结果中常用的“一、±、+、++、+++”,这种资料的研究指标是有序的,并不应该选用 χ^2 检验,而应该选用秩和检验^[3],稍加留意,这类错误也可以在稿件初审时及时发现,见图2。

7、统计学方法:采用 χ^2 检验。

结果: 两组治疗效果差异无统计学意义($P>0.01$)采用桂枝茯苓胶囊显效率达74%,缓解率达20%,总有效率达94%,6例因服药后胃部不适不能坚持服药,无效率为6%;逍遥丸显效率达70%,缓解率达23%,7例因服药后胃部不适未能坚持服药,无效率为7%,总有效率达93%,见表1。

表1 两种药物治疗乳腺增生症疗效比较:

组别	显效	缓解	无效	例数	总有效率
桂枝茯苓胶囊治疗组	74	20	6	100	94%
逍遥丸治疗组	70	23	7	100	93%

图2 某来稿中有序资料错误选取 χ^2 检验作为统计学分析方法

三、 χ^2 检验的基本条件是否满足

χ^2 检验的基本条件除了要求总样本量不小于40之外,还要求任一格子的理论频数 T 不小于1,在审稿时候计算理论频数最小的格子的 T 值即可,即行合计最小值即列合计最小值对应的那一格的 T 值。若这个值大于等于5,则可认为符合基本条件;若该值大于等于1小于5则需要校正,此时应注意论文中是否交待了对使用校正后的 χ^2 值;若该值小于1,则说明该样本不符合 χ^2 检验的基本条件,不能用 χ^2 检验,若论文中使用了 χ^2 检验,也属于统计学错误。理论频数的计算较简单:

某格子的理论频数=(该格子所在行的合计 \times 该格子所在列的合计)/总例数^[2]。

四、检验水准确定是否正确

检验水准(α)的本质是控制假设检验中犯I类错误的概率,即拒绝了原本正确的 H_0 ,通常 H_0 都是假设被检验分组间没有差异,所以可通俗理解为假设检验中的“假阳性”率,检验水准 α 最常用的两个值是0.05和0.01,至于选取哪个,应该是根据研究对象特征和样本量等因素进行的,而且在样本量一定的时候,犯I类错误的概率 α 变小,相应的犯II类错误的概率 β 将变大。例如,如果评价一种新药和某常规药物的疗效,为了提高下结论的谨慎性,不轻易认为新药疗效优于常规药物,可选较小的0.01作为 α 值;而评价一种新的检验方法与某“金标准”对相同标本的检验结果是否一致时,同样为了谨慎,不

轻易认同一种新检验方法等同于金标准,则应该控制更小的 β ,从而选取较大的0.05作为 α 值。

而且,检验水准 α 的大小也并非只能在0.05和0.01中进行选择,在某些情况下,检验水准是需要修正的,例如,多组独立样本的 χ^2 检验中,如果得出“各组不全相同”的结论后要进一步明确是哪两组不同,需要进一步作多组间的两两比较,这时候两两比较需要根据比较的次数来修正检验水准 α ^[2]:

$$\text{校正后的 } \alpha = \text{校正前的 } \alpha / C_R^2$$

R 是组数, C_R^2 就是组数 R 取2的组合数,即所有处理组两两对比有 C_R^2 种情况^[4],例如一个5个处理组的资料,在检验总的概率时选用的 α 为0.05,在作两两对比时,5组共有10种对比情况,则修正后的检验水准取 $\alpha = 0.05/10 = 0.005$ 。如果在审稿中,发现论文进行了多组间的两两比较,但却仍然以修正前的0.05或者0.01为 α ,则也属于统计学的错误,应退给作者重新计算。

值得一提的是,假设检验最终用 P 值与检验水准 α 进行比较来下结论,应该正确理解 P 值的意义,拒绝 H_0 只是说明差异不为零,而 P 值小于 α 也只是说明下这个“有差异”的结论犯错误的概率小于 α ,而 P 值的大小并不反映差异的程度,将 $p < 0.05$ 说成是“差异显著”,将 $p < 0.01$ 说成是“差异非常显著”都是不对的, α 的大小应该是在计算 P 值之前就确定好的,然后通过计算出来的 P 值与其比较来下“差异有统计学意义”或者“差异无统计学意义”的结论^[2]。

总之,医学科论文中统计学运用广泛,作者容易犯的错误也有一定规律性和集中性。作为医学科期刊的编辑,在工作中应该总结一些经验,以便在初审时及时发现及时处理,从而在一定程度上减少不必要的后续审稿流程,提高效率并降低成本。医学统计学庞大而复杂,要做到准确、高效地审核论文中的统计学内容,还需要进一步加强学习。

参考文献:

- [1] 易东. 医学论文中的统计学问题[J]. 重庆医学, 2008, 37(23): 1-8.
- [2] 方积乾. 卫生统计学[M]. 5版. 北京: 人民卫生出版社, 2004: 100-126.
- [3] 杜媛鲲, 姜恒丽, 李玉丁. 医学论文中关于有序资料统计学方法的选用[J]. 临床荟萃, 2008, 23(5): 380.
- [4] 盛骤, 谢式千, 潘承毅. 概率论与数理统计[M]. 北京: 高等教育出版社, 1989: 135-137.

责任编辑: 周仁惠