

P 值大小不等价于差异或相关性大小

郑巧兰, 林燕薇, 王景周

中山大学附属第三医院期刊中心(广州 510630)



【摘要】目的 探讨 P 值与差异或相关性大小的关系。**方法** 固定样本统计量 [均数和标准差、样本率、相关系数、OR (95%CI)], 分别模拟独立样本 t 检验、两组样本率比较的卡方检验、Pearson 相关分析、单因素 Logistic 回归分析的数据, 观察不同样本量下 P 值差异。**结果** 两组样本统计量相同或单组样本统计量一致但样本量不同时, P 值可能大于 0.05、介于 0.001~0.05 或小于 0.001。**结论** P 值受样本量影响大, 其大小不等价于差异或相关性大小。

【关键词】 P 值; t 检验; 卡方检验; Pearson 相关分析; Logistic 回归分析

The size of P value does not equal to the magnitude of difference or correlation

ZHENG Qiaolan, LIN Yanwei, WANG Jingzhou

Journal Center, The Third Affiliated Hospital, Sun Yat-sen University, Guangzhou 510630, China

Corresponding author: WANG Jingzhou, Email: wangjzh53@mail.sysu.edu.cn

【Abstract】Objective To explore the relationship between P value and the size of difference or correlation. **Methods** Fixing sample statistics including mean and standard deviation, sample rate, correlation coefficient and OR (95%CI), then the data of independent sample t -test, chi-square test for comparing two sample rates, Pearson linear correlation analysis and univariate Logistic regression analysis were simulated respectively to observe the difference of P value under different sample size. **Results** When the difference in statistics between two sample groups is the same or the statistics within a single sample group are consistent but the sample size is different, the obtained P values may be >0.05 , between 0.001 and 0.05, or <0.001 . **Conclusion** The P value is greatly affected by the sample size, and its size does not equal to the magnitude of difference or correlation.

【Keywords】 P value; t -test; Chi-square test; Pearson linear correlation analysis; Logistic regression analysis

统计学推断(假设检验)是科学研究尤其是医学研究中的一种常用方法,用于判断样本与总体的一致性。统计学推断普遍依据 P 值, P 值是从一个假设所规定的特定总体中随机抽样,得到的统计量值与实际观测数据相同或更极端(更大或者更小)的概率^[1]。 P 值不代表处理效应的大小

或结果的重要性^[2],但其被曲解和错误使用情况仍较普遍,已公开发表的文章、个别期刊的投稿须知等存在将 $P < 0.05$ 、0.01、0.001 分别解释为差异或相关性显著、非常显著、极显著的现象^[3-7]。目前关于 P 值的研究大多列举 P 值的各类误用情况^[8-9],或者虽然认为不能将 $P < 0.05$ 描述为差

异具有显著性,但未作详细解释。因此,本研究拟从样本量对 P 值影响的角度,探讨 P 值与差异或相关性大小的关系,以期帮助期刊编辑同仁和作者更准确地理解 P 值并解释其意义,提升学术期刊质量。

1 资料与方法

在医药卫生期刊中, t 检验、卡方检验、Pearson 相关分析以及 Logistic 回归分析为使用较多的统计学检验方法,对 P 值的误解也最为常见。本研究通过固定上述 4 种统计方法的主要统计量[均数和标准差、样本率、相关系数、OR 值(95%CI)]模拟数据,观察不同样本量下 P 值差异。

本研究每种统计方法使用 3 份不同样本量的数据进行分析。根据每种统计方法适用的数据分布情况,设定统计量值和样本量,采用随机数函数生成研究数据。以 t 检验为例,设定两组样本的均数、标准差和样本量,采用正态分布随机数函数分别随机生成两组样本的第一份数据,即模拟 1;为确保统计量值完全一致,复制 10 次第一份数据得到第

二份数据,即模拟 2;复制 20 次第一份数据得到第三份数据,即模拟 3。研究中,对统计量值和样本量不断调整,使每种统计方法 3 次模拟的 P 值分别 > 0.05 、 $0.001 \sim 0.05$ 、 < 0.001 。

采用 SPSS 20.0 软件进行模拟数据生成和统计检验,采用双侧检验结果。

2 结果

2.1 以独立样本 t 检验为例

设两组样本的体质量指数(body mass index, BMI)分别为 $(20.3 \pm 1.5) \text{ kg} \cdot \text{m}^{-2}$ 、 $(20.0 \pm 1.5) \text{ kg} \cdot \text{m}^{-2}$,当样本量均为 30 时,两组均数比较差异无统计学意义($P=0.444$);当样本量均为 300 时, $P=0.014$;当样本量均为 600 时, $P < 0.001$,见表 1。

2.2 以两组样本率比较的卡方检验为例

设两种治疗方法的有效率分别为 0.42、0.50,当两组样本量均为 50 时,两组有效率比较差异无统计学意义($P=0.422$);两组样本量均为 500 时, $P=0.011$;两组样本量均为 1 000 时, $P < 0.001$,见表 2。

表1 均数和标准差相同、样本量不同的两组数据 t 检验统计分析结果

Table 1. Statistical results of t -test for two groups with constant mean and standard deviation under different sample size

数据	$n_1^{1)}$	$n_2^{2)}$	$\bar{X}_1^{3)}$	$\bar{X}_2^{4)}$	$\bar{X}_1 - \bar{X}_2$	$S_1^{5)}$	$S_2^{6)}$	t 值	ν	P 值
模拟1	30	30	20.3	20.0	0.3	1.5	1.5	0.773	58	0.444
模拟2	300	300	20.3	20.0	0.3	1.5	1.5	2.476	598	0.014
模拟3	600	600	20.3	20.0	0.3	1.5	1.5	3.505	1 198	< 0.001

注: ^{1)、2)}分别表示两组样本的样本量; ^{3)、4)}分别表示两组样本的均数; ^{5)、6)}分别表示两组样本的标准差; ν 为自由度。

表2 有效率相同、样本量不同的两组数据卡方检验统计分析结果

Table 2. Statistical results of chi-square test for two groups with constant effective rate under different sample size

数据	$n_1^{1)}$	$n_2^{2)}$	$N_1^{3)}$	$N_2^{4)}$	$p_1^{5)}$	$p_2^{6)}$	$p_2 - p_1$	χ^2 值	ν	P 值
模拟1	21	25	50	50	0.42	0.50	0.08	0.644	1	0.422
模拟2	210	250	500	500	0.42	0.50	0.08	6.441	1	0.011
模拟3	420	500	1 000	1 000	0.42	0.50	0.08	12.882	1	< 0.001

注: ^{1)、2)}分别表示两组样本的治疗有效例数; ^{3)、4)}分别表示两组样本的总例数; ^{5)、6)}分别表示两组样本的治疗有效率; ν 为自由度。

2.3 以 Pearson 相关分析为例

两个变量相关系数为 0.238,样本量均为 10 时, $P=0.508$;样本量扩大至 100 时, $P=0.017$;样本量为 1 000 时, $P < 0.001$,见表 3。

2.4 以单因素 Logistic 回归分析为例

单因素 Logistic 回归分析中, $\beta=0.811$,OR 值为 2.250,当两组样本量均为 10 时,OR 值 95%CI 为 $(0.376, 13.465)$, P 值为 0.374;样本

量增加到每组 100 时, OR 值不变, 95%CI 缩窄为 (1.278, 3.962), P 值为 0.005; 样本量为每

组 200 时, OR 值不变, 95%CI 为 (1.508, 3.357), $P < 0.001$, 见表 4。

表3 相关系数相同、样本量不同的两个变量Pearson相关分析结果

Table 3. Results of Pearson correlation analysis for two variables with constant correlation coefficients under different sample size

数据	$n_1^{1)}$	$n_2^{2)}$	r 值	ν	P 值
模拟1	10	10	0.238	18	0.508
模拟2	100	100	0.238	198	0.017
模拟3	1 000	1 000	0.238	1 998	<0.001

注: ¹⁾、²⁾分别表示两个变量的样本量; ν 为自由度。

表4 回归系数相同、样本量不同的两组数据单因素Logistic回归分析结果

Table 4. Results of binary Logistic regression analysis for two groups with constant regression coefficient under different sample size

数据	$n_1^{1)}$	$n_2^{2)}$	β	OR (95%CI)	P 值
模拟1	10	10	0.811	2.250 (0.376, 13.465)	0.374
模拟2	100	100	0.811	2.250 (1.278, 3.962)	0.005
模拟3	200	200	0.811	2.250 (1.508, 3.357)	<0.001

注: ¹⁾、²⁾分别表示两组的样本量。

3 讨论

了解检验功效 (power) 有助于正确理解 P 值。在统计学中, 检验功效是指当两个总体参数间存在差异时, 所使用的统计检验能够发现这种差异的概率。影响检验功效的因素通常有: 总体参数间差异越大, 检验功效越大; 个体差异 (标准差) 越小, 检验功效越大; 样本含量越大, 检验功效越大; 检验水准 (α) 越大, 检验功效越大^[10]。检验功效与 P 值呈负相关^[11]。因此, 在其他条件相同的情况下, 样本量越大, 检验功效越大, 即能够发现这种差异的概率越大, 而 P 值越小。

本研究验证了 P 值受样本量影响较大: 将每种统计方法 3 次模拟的统计量值固定, 即从统计描述上看数据差异或相关性一致, 但由于样本量不同, 得到的 P 值差异很大, 甚至得到完全不同的结果。如果将更小 P 值的结果解释为差异更大或者显著, 可能造成疑惑。例如, 在两组独立样本 t 检验中, 虽然 3 种样本量得到不同的 P 值, 但是 3 种样本下两组的均数差值均为 0.3, 差异从统计描述上看一致。而且, 虽然第 2、3 次模拟数据显示 P 值小于 0.05, 但是在临床上相差 0.3 可能没有专业意义, 甚至两组均在正常的 BMI 范围内。同样地, 在两组有效率比较的卡方检验中, 3 种不同样本量情况下两组

的治疗有效率一致, 样本量增加后 P 值变小, 并非代表差异更加显著。正确的理解应该是以目前的情况, 出现零假设正确, 即无差异的概率更低, 更有理由认为两组的有效率有差异。在 Pearson 相关分析中, 3 次模拟数据的相关系数均为 0.238, 第 1 次的模拟数据因为 P 值大于 0.05, 提示两组数据无线性相关, 但在第 2、3 次的模拟数据中 P 值均小于 0.05, 而使不少作者认为显著相关^[5]。实际上, 相关系数的假设检验仅能说明相关关系是否有统计学意义, 不能说明相关关系是否密切^[12]。有统计学家认为, $r \geq 0.7$ 为强相关, 在 0.3 以下为弱相关甚至无相关^[13]。在单因素 Logistic 回归分析中, 当 OR 值不变而样本量增加时, P 值变小, OR 值的 95%CI 缩窄, 这是由于 95%CI 与标准误相关, 而标准误受样本量的影响, 更窄的 95%CI 意味着更稳固的结果、更高的可信度^[10, 14]。

新英格兰医学杂志 2019 年的作者指南指出, P 值没有提供效应大小或者关联强弱的信息, 其投稿指南要求 P 值大于 0.01 时保留 2 位小数; 如果 P 值在 0.001~0.01 之间, 则保留 3 位小数; 如果 P 值小于 0.001, 用 $P < 0.001$ 表示, 同时应提供效应值、关联度或其他感兴趣结果指标的置信区间^[15-16]。在国内, 由于学术不端事件频发, 需要进行结果核对, 因此, 期刊编辑应让作者提供

具体 P 值及其他统计量值^[17-18]。中国高校医学期刊论文统计报告规范编写组近期推出了《中国高校医学期刊论文统计报告推荐清单(2023 版)》，指出“大于 0.001 的 P 值应保留 3 位小数，小于 0.001 的 P 值应报告 $P < 0.001$ ”，但并未对 P 值做出更详细解释或设置更多规范^[19]。本研究对 P 值的解释提出以下建议：首先，统计学方法部分不应将 $P < 0.05$ 、0.01、0.001 表达为差异或相关性显著、非常显著、极显著，仅需表达检验水准为 0.05（或其他）或者 $P < 0.05$ 为差异有统计学意义即可。其次，结果部分应避免出现带有评论意义的“显著”“明显”等词汇，尤其是不应根据 P 值大小判断差异是否显著、非常显著、极显著。英文写作中的“statistical significance”翻译为中文是“统计学显著性”，其代表的是数据差异在统计学上是否显著，不代表专业上的显著性。但是不少论文作者混淆概念，将统计学显著性等价于差异显著性，文字上表述为“两组比较具有显著差异”或“显著高于、明显高于”等，降低了文章的严谨性，容易对读者造成严重误导。本研究建议当 P 值小于设定的检验水准时，不论其大小，仅表达差异具有统计学意义并附上对应的 P 值即可。另外，讨论部分也应慎重表述差异显著的结论，是否显著应根据效应值大小及其置信区间并结合临床专业意义进行判断；当 P 值大于 0.05 时，还应结合样本量及检验功效讨论。

值得注意的是，不能唯 P 值论，结果报告需要更多地结合效应量和置信区间^[20]，贝叶斯等方法也越来越多地被推荐；同时，统计检验中注意多重比较校正，以减少错误发现的可能性；且统计学意义并不代表专业意义，研究者还应考虑研究结果的专业意义。

P 值相关问题一方面反映出国内统计学教育的不足，是统计学教学过程中相对薄弱且需要重点关注的环节^[21]；另一方面也体现了继续教育的欠缺，以期刊编辑工作为例，编辑的统计学素养可强化作者的统计学意识，提升科研人员的统计学水平，其统计学掌握水平与论文综合质量密切相关，国内学术期刊编辑继续教育存在内容同质化、培训方式单一、培训对象宽泛等特征^[22-24]。学术期刊编辑尤其是医学期刊编辑亟须提高统计学素养，期刊编辑继续教育相关组织机构可以安排医药卫生期刊专场，邀请统计学专家进行授课；

对于论文作者，编辑部可以通过完善稿约细则，增强作者统计报告规范意识，引导其重视统计学规范，正确使用 P 值解释相关结果^[25]。

综上， P 值虽然饱受争议，但目前仍为医学科研论文不可忽略的关键词，正确使用并解释 P 值对于科学研究的可靠性、结论推断和决策制定具有重要意义。本研究通过模拟数据分析，展示了样本量对 P 值大小的影响，说明了 P 值大小不等同于差异或相关性大小的原因，提出正确使用和解释 P 值的建议，以期帮助编辑同仁和作者准确理解 P 值并正确解释统计分析结果，提高学术期刊论文质量。

参考文献

- 1 Wasserstein RL, Lazar NA. The ASA statement on p -values: context, process, and purpose[J]. *Am Stat*, 2016, 70(2): 129-133. DOI: [10.1080/00031305.2016.1154108](https://doi.org/10.1080/00031305.2016.1154108).
- 2 周敏林. 临床试验中 P 值的意义及结果解读[J]. *肾脏病与透析肾移植杂志*, 2017, 26(2): 170-173, 183. [Zhou ML. Interpretation of P value in clinical trials[J]. *Chinese Journal of Nephrology, Dialysis & Transplantation*, 2017, 26(2): 170-173, 183.] DOI: [10.3969/cndt.j.issn.1006-298X.2017.02.015](https://doi.org/10.3969/cndt.j.issn.1006-298X.2017.02.015).
- 3 杨灵敏. 分析超声诊断乳腺癌腋窝淋巴结转移的影像学表现及其临床效果[J]. *中国医药指南*, 2023, 21(9): 106-108. [Yang LM. To analyze the imaging features and clinical effects of ultrasound diagnosis of axillary lymph node metastasis in breast cancer[J]. *Guide of China Medicine*, 2023, 21(9): 106-108.] DOI: [10.15912/j.cnki.gocm.2023.09.012](https://doi.org/10.15912/j.cnki.gocm.2023.09.012).
- 4 王晓玲. 益肺健脾汤与三伏灸贴联合治疗老年慢性支气管炎的临床效果[J]. *中医临床研究*, 2021, 13(25): 101-103. [Wang XL. Clinical effects of the Yifei Jianpi decoction plus Sanfu moxibustion plaster on chronic bronchitis in the elderly[J]. *Clinical Journal of Chinese Medicine*, 2021, 13(25): 101-103.] DOI: [10.3969/j.issn.1674-7860.2021.25.030](https://doi.org/10.3969/j.issn.1674-7860.2021.25.030).
- 5 姚梦, 唐东辉, 白爽. 综合干预对肥胖男性青少年血管生成素样蛋白 2 和血管内皮功能的影响[J]. *中国学校卫生*, 2021, 42(12): 1888-1891. [Yao M, Tang DH, Bai S. Effect of comprehensive intervention on ANGPTL2 and vascular endothelial function in obese male adolescents[J]. *Chinese Journal of School Health*, 2021, 42(12): 1888-

- 1891.] DOI: [10.16835/j.cnki.1000-9817.2021.12.031](https://doi.org/10.16835/j.cnki.1000-9817.2021.12.031).
- 6 中国新药杂志编辑部.《中国新药杂志》投稿须知[EB/OL]. (2023-03-03) [2023-12-01]. [Editorial of Chinese Journal of New Drugs. Guideline of Chinese Journal of New Drugs [EB/OL]. (2023-03-03) [2023-12-01].] <http://www.newdrug.cn/CN/column/column9.shtml>.
- 7 万里, 孙劲楠, 丁佐奇. 封面论文和非封面论文的学术影响力对比研究——基于药理学类核心期刊的实证研究[J]. 科技与出版, 2022(11): 145-151. [Wan L, Sun JN, Ding ZQ. Comparative study on academic influence of cover papers and non-cover papers—an empirical study based on pharmaceutical core journals[J]. Science-Technology and Publication, 2022(11): 145-151.] DOI: [10.16510/j.cnki.kjyeb.20221012.009](https://doi.org/10.16510/j.cnki.kjyeb.20221012.009).
- 8 吴艳妮, 周春兰, 江霞, 等. 国内护理学统计源期刊论文中报告精确 P 值常见错误: $P=0.000$ [J]. 编辑学报, 2016, 28(2): 133-134. [Wu YN, Zhou CL, Jiang X, et al. Common errors in reports on accurate P value in papers from domestic nursing journals[J]. Acta Editologica, 2016, 28(2): 133-134.] DOI: [10.16811/j.cnki.1001-4314.2016.02.010](https://doi.org/10.16811/j.cnki.1001-4314.2016.02.010).
- 9 汪媛, 张爽爽, 应倩, 等. 国内肿瘤学期刊论文中 P 值为零的误用分析[J]. 编辑学报, 2017, 29(1): 46-47. [Wang Y, Zhang SS, Ying Q, et al. Misuse of the P value equal to zero in papers of Chinese oncology journals[J]. Acta Editologica, 2017, 29(1): 46-47.] DOI: [10.16811/j.cnki.1001-4314.2017.01.015](https://doi.org/10.16811/j.cnki.1001-4314.2017.01.015).
- 10 方积乾. 卫生统计学(第6版)[M]. 北京: 人民卫生出版社, 2008. [Fang JQ. Medical Statistic (6th edition) [M]. Beijing: People's Medical Publishing House, 2008].
- 11 钱俊, 陈平雁. 假设检验中计算观察检验效能的意义的探讨[J]. 中国卫生统计, 2005, 22(3): 133-137. [Qian J, Chen PY. What's the meaning of observed power in hypothesis testing[J]. Chinese Journal of Health Statistics, 2005, 22(3): 133-137.] DOI: [10.3969/j.issn.1002-3674.2005.03.002](https://doi.org/10.3969/j.issn.1002-3674.2005.03.002).
- 12 王雅洁, 李亚琴. 医学科技论文中常见统计学问题分析及建议[J]. 护理研究, 2021, 35(10): 1757-1760. [Wang YJ, Li YQ. Analysis and suggestions on common statistical problems in medical scientific papers[J]. Chinese Nursing Research, 2021, 35(10): 1757-1760.] DOI: [10.12102/j.issn.1009-6493.2021.10.010](https://doi.org/10.12102/j.issn.1009-6493.2021.10.010).
- 13 Rumsey DJ. Statistics for dummies[M]. Hoboken NJ: Wiley Publishing, 2011.
- 14 冯国双. 统计学中 P 值的含义及其应用中的误解[J]. 中华全科医师杂志, 2017, 16(1): 83-84. [Feng GS. The interpretation and misunderstanding about P value in statistics[J]. Chinese Journal of General Practitioners, 2017, 16(1): 83-84.] DOI: [10.3760/cma.j.issn.1671-7368.2017.01.021](https://doi.org/10.3760/cma.j.issn.1671-7368.2017.01.021).
- 15 Harrington D, D'Agostino RB Sr, Gatsonis C, et al. New guidelines for statistical reporting in the journal[J]. N Engl J Med, 2019, 381(3): 285-286. DOI: [10.1056/NEJMe1906559](https://doi.org/10.1056/NEJMe1906559).
- 16 The New England Journal of Medicine. Statistical Reporting Guidelines[EB/OL]. [2023-12-09]. <https://www.nejm.org/author-center/new-manuscripts>.
- 17 相丹凤, 高永, 周英智. 医学论文中成组 t 检验 P 值错误及其原因分析[J]. 中国科技期刊研究, 2018, 29(12): 1224-1228. [Xiang DF, Gao Y, Zhou YZ. P -value errors in two-sample t -test of medical papers and reason analysis[J]. Chinese Journal of Scientific and Technical Periodicals, 2018, 29(12): 1224-1228.] DOI: [10.11946/cjstp.201807300679](https://doi.org/10.11946/cjstp.201807300679).
- 18 郑巧兰, 林燕薇, 洪悦民, 等. 科技期刊编辑如何利用统计知识筛查学术不端行为[J]. 编辑学报, 2021, 33(1): 74-77. [Zheng QL, Lin YW, Hong YM, et al. How can editors of scientific journals use statistical method to screen academic misconduct[J]. Acta Editologica, 2021, 33(1): 74-77.] DOI: [10.16811/j.cnki.1001-4314.2021.01.017](https://doi.org/10.16811/j.cnki.1001-4314.2021.01.017).
- 19 中国高校医学期刊论文统计报告规范编写组. 中国高校医学期刊论文统计报告推荐清单(2023版)[J]. 郑州大学学报(医学版), 2023, 58(6): 875-882. [Chinese university medical journal paper statistical report standard compilation group. Recommended list of statistical reports of medical journals in colleges and universities in China (2023 edition)[J]. Journal of Zhengzhou University (Medical Sciences), 2023, 58(6): 875-882.] DOI: [10.13705/j.issn.1671-6825.2023.10.004](https://doi.org/10.13705/j.issn.1671-6825.2023.10.004).
- 20 黄申, 蒋青青, 王世琦, 等. P 值和置信区间: 联系与区别、误用与争论[J]. 数理医药学杂志, 2023, 36(1): 3-8. [Huang S, Jiang QQ, Wang SQ, et al. P -value and confidence interval: connection and difference, misuse and argument[J]. Journal of Mathematical Medicine, 2023, 36(1): 3-8.] DOI: [10.12173/j.issn.1004-4337.202212021](https://doi.org/10.12173/j.issn.1004-4337.202212021).
- 21 彭佳师. P 值争论对生物统计学课程教学的影响

- [J]. 生物学杂志, 2020, 37(6): 120–122, 129. [Peng JS. *P*-value controversy and its impact on biostatistics teaching[J]. Journal of Biology, 2020, 37(6): 120–122, 129.] DOI: [10.3969/j.issn.2095-1736.2020.06.120](https://doi.org/10.3969/j.issn.2095-1736.2020.06.120).
- 22 李晓炜, 王春燕, 王希营, 等. 医学期刊青年编辑统计学素养的影响因素和提升策略 [J]. 编辑学报, 2020, 32(2): 225–227. [Li XW, Wang CY, Wang XY, et al. Influence factors and promotion strategies of statistical accomplishment for young editors of medical journals[J]. Acta Editologica, 2020, 32(2): 225–227.] DOI: [10.16811/j.cnki.1001-4314.2020.02.028](https://doi.org/10.16811/j.cnki.1001-4314.2020.02.028).
- 23 李玥, 栾嘉, 邓强庭, 等. 编辑因素对我国中文医学期刊学术质量影响的实证研究 [J]. 中国科技期刊研究, 2023, 34(3): 288–296. [Li Y, Luan J, Deng QT, et al. Empirical study on the influence of factors related to editors on the academic quality of Chinese medical journals[J]. Chinese Journal of Scientific and Technical Periodicals, 2023, 34(3): 288–296.] DOI: [10.11946/cjstp.202210100766](https://doi.org/10.11946/cjstp.202210100766).
- 24 李伟. 2016—2022 年中国科技期刊编辑继续教育培训的特征分析 [J]. 中国科技期刊研究, 2023, 34(5): 647–652. [Li W. Characteristics of continuing education training for Chinese scientific journal editors from 2016 to 2022[J]. Chinese Journal of Scientific and Technical Periodicals, 2023, 34(5): 647–652.] DOI: [10.11946/cjstp.202303090147](https://doi.org/10.11946/cjstp.202303090147).
- 25 易耀森. 医学期刊稿约中统计报告规范的现状及完善建议 [J]. 编辑学报, 2022, 34(5): 510–514. [Yi YS. Constructive suggestions on statistical reporting standards of author guidelines in medical journals[J]. Acta Editologica, 2022, 34(5): 510–514.] DOI: [10.16811/j.cnki.1001-4314.2022.05.008](https://doi.org/10.16811/j.cnki.1001-4314.2022.05.008).

收稿日期: 2023 年 01 月 21 日 修回日期: 2024 年 02 月 20 日
本文编辑: 张 苗 黄 笛

引用本文: 郑巧兰, 林燕薇, 王景周. *P*值大小不等价于差异或相关性大小[J]. 数理医药学杂志, 2024, 37(3): 158–163. DOI: [10.12173/j.issn.1004-4337.202401131](https://doi.org/10.12173/j.issn.1004-4337.202401131).
Zheng QL, Lin YW, Wang JZ. The size of *P* value does not equal to the magnitude of difference or correlation[J]. Journal of Mathematical Medicine, 2024, 37(3): 158–163. DOI: [10.12173/j.issn.1004-4337.202401131](https://doi.org/10.12173/j.issn.1004-4337.202401131).