

ASA 主席关于统计显著性和可重复性的专责小组声明

作者: YOAV BENJAMINI¹, RICHARD D. DE VEAUX², BRADLEY EFRON³, SCOTT EVANS⁴, MARK GLICKMAN⁵, BARRY I. GRAUBARD⁶, XUMING HE⁷, XIAO-LI MENG⁸, NANCY REID⁹, STEPHEN M. STIGLER¹⁰, STEPHEN B. VARDEMAN¹¹, CHRISTOPHER K. WIKLE¹², TOMMY WRIGHT¹³, LINDA J. YOUNG¹⁴ AND KAREN KAFADAR¹⁵

在过去的十年中, 科学界对研究结果的可重复性越来越关注. 可重复性的一个重要方面是使用统计方法来构造结论的框架. 2019 年, 美国统计协会 (ASA) 主席成立了一个专责小组, 以解决人们担心 *The American Statistician* (ASA 期刊) 2019 年的社论可能被错误地解释为 ASA 的官方政策. (2019 年的社论建议在统计分析中取消使用 ' $p < 0.05$ ' 和 "统计显著性") 本文件是 ASA 邀请我们公布的专责小组的声明. 其目的有两个: 首先, 经过恰当应用及解释后, P 值和显著性的使用是不应放弃的重要工具; 其次, 简要阐明一些可能对科学界有用的统计推断的一些健全原则.

P 值是有效的统计量度, 为传达定量结果中固有的不确定性提供了方便的惯例. 事实上, P 值和显著性检验是统计文献中研究最多和最深入理解的统计方法之一. 它们是通过适当的应用而推进科学的重要工具.

围绕统计显著性的很多争议可以通过更好地认识不确定性、多变性、多重性和可重复性来消除. 以下一般原则是适当使用 P 值和报告统计显著性并更广泛地用于良好统计实践的基础.

捕捉与统计结论相关的不确定性至关重要. 不同的不确定性度量可以相互补充, 没有任何一项度量能服务于所有目的. 统计结论所描述的变异来源应在科学文章和报告中加以描述. 在可能的情况下, 还应识别那些尚未描述的变异来源.

处理可重复性和不确定性是统计科学的核心. 如果研究结果能够用新数据在进一步的研究中得到验证, 则研究结果是可重复的. 撇开欺诈的可能性不谈, 可重复性出问题的重要来源包括研究设计和实施不佳、数据不足、缺乏对模型选择的关注而没有充分了解该选择的含义、对分析和计算程序的描述不充分、以及对报告结果的选择. 即使是在报道的部分中突出一些有说服力的结果, 选择性地报道都可能导致对证据的扭曲看法. 在某些情况下, 可以通过调整多重性来缓解此问题. 不确定性的控制和解释应开始于研究和测量过程的设计, 并持续到分析的每个阶段, 直至报告结果. 即使在精心设计、精心执行的研究中, 固有的不确定性依然存在, 统计分析也应适当解释这种不确定性.

统计学的理论基础为应对不确定性提供了若干一般策略. P 值、置信区间和预测区间通常与频率派方法相关. 贝叶斯因子、后验概率分布和可信区间通常用于贝叶斯方法. 这些是用于反映不确定性的许多统计方法中的一部分.

当需要做决策时, 阈值是有帮助的. 尽管 P 值本身提供了有价值的信息, 但将 P 值与显著性水平进行匹配可能很有用. P 值和统计显著性应理解为对与抽样变化相关的观测或影响的评估, 而不一定理解为具有实际显著性的度量指标. 如果认为有必要将阈值视为决策的一部分, 则应根据研究目标明确定义阈值, 并考虑错误决策的后果. 作为惯例的阈值 (显著性水平—译者) 标准应依照学科和分析目标的差异而变化.

总之, P 值和显著性检验在正确应用和解释时会增加从数据中得出的结论的严谨性. 分析数据和总结结果往往比通常人们认识的要复杂得多. 虽然所有科学方法都有局限性, 但正确应用统计方法对于解释数据分析的结果和提高科学结果的可重复性至关重要.

“那些自称让事实和数字为自己说话、同时又将他在对事实和数字做选择和分组时所扮演的角色 (也许是无意识地) 隐藏在幕后的人是所有理论家中最鲁莽和最好诈的.”

(Alfred Marshall, 1885)

补充材料

对“ASA 主席关于统计显著性和可重复性的专责小组声明”的补充 (DOI: 10.1214/21-AOAS1501SUPP;.pdf). 专责小组成员的简介.

¹Department of Statistics and Operations Research, Tel Aviv University, ybenja@tauex.tau.ac.il

²Department of Mathematics and Statistics, Williams College, deveaux@williams.edu

³Department of Statistics and Department of Biomedical Data Sciences, Stanford University, brad@stat.stanford.edu

⁴Department of Biostatistics & Bioinformatics, George Washington University, sevens@bsc.gwu.edu

⁵Department of Statistics, Harvard University, glickman@fas.harvard.edu

⁶Biostatistics Branch, National Cancer Institute, barry.graubard@nih.gov

⁷(Co-chair) Department of Statistics, University of Michigan, xmhe@umich.edu

⁸Department of Statistics, Harvard University, meng@stat.harvard.edu

⁹Department of Statistics, University of Toronto, reid@utstat.utoronto.ca

¹⁰Department of Statistics, University of Chicago, stigler@uchicago.edu

¹¹Department of Statistics and Department of Industrial & Manufacturing Systems Engineering, Iowa State University, vardeman@iastate.edu

¹²Department of Statistics, University of Missouri, wikle@missouri.edu

¹³Center for Statistical Research and Methodology, United States Bureau of the Census, tommy.wright@census.gov

¹⁴(Co-chair), Research & Development, National Agricultural Statistics Service, linda.j.young@usda.gov

¹⁵(Ex-officio) Department of Statistics, University of Virginia, kkafadar@virginia.edu