



科学事实与统计思维

文
／
程
开
明

科学常识与素养是公民素质的重要组成部分，反映一个国家或地区的软实力，从根本上制约着创新能力的提升和社会经济发展。英国剑桥大学教授苏斯伦德 (William J. Sutherland) 等在《自然》杂志 2013 年第 503 卷 21 期上提出“解读科学观点时应该知道的 20 个事实”，阅后发现其中提到的科学事实都与统计思维有关。现代科学研究中统计学是最重要的工具之一，英国著名生物学家高尔顿曾说过：“统计学具有处理复杂问题的非凡能力，当科学的探索者在前进的过程中荆棘载途时，唯有统计学可以帮助他们打开一条通道。”当然，统计也很容易被误用或滥用，《自然》杂志 2014 年第 506 卷 13 期刊发评论《数字危机》，宣称“很多研究者缺乏统计学基本知识，普遍存在统计误用现象，在一些基础科学领域这种情况十分严重”。所以，运用科学研究结论辅助现实决策时，须具备良好的统计思维，才能对科学结论保持清晰认识，更准确地解读结论背后的科学真相。

一、人人应该知道的 20 个科学事实

苏斯伦德等教授认为将科学结论应用于政策分析，目前仍存在许多严峻的问题，当务之急是要让政策制定者认识到科学那不完美的本质，具备解读科学结论的能力。鉴于此，公务员、政治人物、政策顾问、记者以及其他与科学打交道的非专业人士，在根据相关研究结论制定政策时都需要知道 20 个基本事实。具体如下：

1. 事实差异和偶然都会引起变化。现实世界的变化是不可预测的，而科学竭力探究是什么因素造成这些变化，譬如为什么现在比过去十年更热？等等。对于这些现象的科学解释，最大挑战是从无数可能因素中梳理出我们感兴趣的过程所起到的作用。

2. 没有绝对准确的测量。事实上，所有测量都存在误差，每一次重复试验都可能得到不同的结果。在某些情况下，与实际变化相比测量误差也许会很大。提出结论时应该给出结果的误差范围，

以避免对精确度的不合理判断。

3. 偏倚是很普遍的。实验设计或测量装置可能对结果造成影响，容易在某一特定方向上产生非典型结果。样本的构成不同，可能得到不同的结果。那些“统计上显著”的结果更可能被报道和出版，只看文献易让人产生错觉——问题的严重性或方案的有效性往往被夸大。

4. 样本量通常越大越好。大样本得到的观测结果往往比小样本更稳定，这一点对那些变量多、容易产生测量

误差的复杂体系尤为重要。譬如，药物测试中为了更可靠、准确地估计药物的平均功效，一个有数以万计样本量的实验要比一个只有数百个样本量的实验好得多。

5. 相关不一定意味着因果。相关关系的得出有时候是出于偶然，看似存在的相关关系可能是由复杂的或潜在的第三方因素所导致。譬如，生态学家曾一度认为有毒的藻类杀死了一条河流入海口处的鱼，但后来发现是因为鱼死了所以藻类繁盛起来。

6. 均值回归效应可能造成误导。单次测量中有一部分极端数据是由于偶然或误差造成的，另一次测量中数据可能就没有那么极端。比如，在经常发生车祸的地方放置一个测速相机，随后事故率的下降并不能归因于放了测速相机，只是事故率回归到正常水平而已。

7. 超越数据范围的推断存在风险。某一范围得到的模型如果超出这个范围可能就不再适用，譬如如果现在的气候变化速度比现存物种在进化史中所经历过的任何时期都要快，或是出现一个全新的极端天气系统时，评价生态系统对于气候变化的反馈就十分困难。

8. 注意基础比率谬误。一个不完美的检验到底有多准？不仅与检验的好坏有关，还与检测状态本身发生的概率有关。譬如，一个人做了有 99% 准确度的血液测试来检测一种罕见疾病，结果虽呈阳性，但受基础比率的影响他健康的可能性比生病的可能性更大。

9. 对照组很重要。对照组可以帮助科学家确保没有额外的因素在干扰结果，没有对照组就难以知道实验处理究竟对结果有没有影响。有时人们在药物测试中表现出阳性，可能受到环境、提供测试的人甚至药丸颜色的影响，此时设置对照组（安慰剂组）非常重要。

10. 随机化能够避免偏倚。实验应尽可能随机地采集样本。例如，直接比较有参与健康计划的家庭和没有参与健康计划的家庭中孩子的学习成绩，很容易带上偏倚（受教育程度较高的家庭更可能参与健康计划），好的实验设计

应当随机地选择参与或不参与的家长。

11. 寻求“重复”而非“伪重复”。在多个研究、独立总体中重复出现的结论更为可靠，如果直接把在某一总体中得到的实验结果推广到其他不具有相同特点的总体，就犯了“伪重复”的错误，伪重复导致研究者对结果产生虚假信心。

12. 科学家也是人。科学家也从工作中获得一些既得利益，包括地位、研究经费甚至是直接的经济利益，这可能导致研究结论的刻意选择与夸大。同行评价并非绝对可靠，期刊编辑可能更支持积极、富有新闻价值的结果。多渠道、独立来源的数据及可重复结果才更可信。

13. 显著性很重要。统计显著性表示一个事件出于偶然而发生的概率，科学家习惯将 $P < 0.05$ 的情况称为显著。比如一项实验中实验组与对照组的差异显著性是 $P=0.01$ ，表示有百分之一的可能性是：实验处理其实没有效果，只是偶然因素导致了实验组和对照组的差异。

14. 不显著不代表没效果。统计上不显著 ($P>0.05$) 并不代表真正的无效，只意味着影响效应未被检测到而已。譬如用基因改造的抗虫棉和抗虫马铃薯做一组实验显示，这些作物对传粉者等益虫不存在不利影响，事实可能是实验的样本量不够大，不足以检测到影响效应。

15. “效应量”很重要。显著性可以衡量差异是“真”还是“假”，但如果差异是“真的”，还需进一步考察具体效应有多大？一项多次重复的实验也许得到统计上显著、效应量很小的结果，这种结果可能并没有实际意义。

16. “关联性”限制结论的推广。科学研究结论能否应用于实际问题，取决于研究条件和实际情况的相似程度。譬如，如果将从动物实验或实验室实验中得到的结果推广到人类的时候，就存在较大的局限性。

17. 感觉会影响风险感知。很多因素都对人类的风险感知造成不同程度的影响，包括事件的罕见性、人们自以为对事件的掌控程度、结果的不利

影响等。比如，美国人往往严重低估在家携带手枪的危险（缩小 100 倍），而严重高估住在核反应堆旁边的危险（扩大 10 倍）。

18. 相关性会改变风险。计算独立事件的后果较为容易，像极潮、强降雨和关键员工的缺席等，但如果这些事件相互关联（风暴导致高水位，而强降雨导致关键员工的缺席），它们共同发生的概率就比预期大得多。

19. 数据可以选择性呈现。有时为了支持自己的观点，实验者会选择对预期结果有利的证据。譬如，一项研究认为怀孕时的酸奶摄入量和后代患哮喘之间显著相关，解读结论首先要知道研究者是预先打算验证这一假说，还是从一大堆数据中偶然发现这一相关性的。

20. 极端测量值可能引起误导。由于个体能力差异、取样、偏见、测量误差等因素的影响，所有测量结果都具有可变性。在解释研究结果的时候，除个体差异之外的因素常被忽略，如果讨论的是极端结果，仅比较极端值和平均值的偏离幅度，就可能带来严重的误导。

上述 20 个科学事实，基本上都与统计常识有关，可见在开展科学研究、理解研究结论以及利用科学结论指导现实决策的过程中，需要具备一定的统计思维，如此才能避免统计陷阱和统计谬误。正如 C.R. 劳教授所讲：“对统计学的一知半解，常常造成不必要的上当受骗；对统计学的一概排斥，往往造成不必要的愚昧无知。”

二、统计思维对于理解科学事实的重要性

作为科学的认识活动，人对客观世界有意识的介入大致包括自然研究和人文社科研究两种基本方式，自然研究的认识对象是自然界的万事万物，基本认识手段就是观察，人文社科研究的认识对象是人世间的万事万物，认识手段从观察扩展至调查。而无论是观察还是调查，确保研究结论的科学性，都须依据基本的统计原理。

科学研究的特征之一是理论体系

的逻辑性，从假设到理论再到推论，一定要有逻辑性，而且理论必须能够解释现实，即理论与现实要有一致性。检验理论与现实之间的一致性，通常是统计学的工作，统计学家在其中扮演着重要角色，包括帮助设计收集数据的方法，提供数据特征的描述方法，以及利用样本数据对总体特征做出估计、检验和预测。

现代社会从信息不足转变为信息泛滥，信息匮乏的危机让位给信息甄别的困难，如此背景下科学方法成为每个人的必修课。在日益依赖数据的今天，树立正确的统计思维，才能有效地开展数据处理与分析。逻辑思维往往隐含一定的前提条件，即使逻辑思维过程非常正确，如果前提条件不满足或者错了，得到的结论可能与现实不符，所以需要利用统计方法来对研究结论进行经验验证。当今世界正步入信息爆炸的大数据时代，统计越显重要，验证了英国科幻小说作家 H.G. 威尔斯的预言：“统计思维总有一天会像读写一样，成为一个有效率公民的必备能力。”

从随机性中寻找规律性，是统计的基本思想，也是统计的魅力所在。统计学被广泛应用于各门学科之中，从自然科学到人文社会科学，甚至是工商业及政府的情报决策。作为认识自然、社会的工具和手段，统计研究客观现象的数量关系，帮助政策制定者理解科研证据对决策的作用。正如现代统计学的奠基人费歇尔所讲：“给 20 世纪带来了人类进步的独特方面是统计学，统计学的普遍存在以及在开拓新知识领域方面的应用已远远超过 20 世纪内的任何技术或科学发明。”（李金昌，2009）

马寅初曾说：“学者不能离开统计而究学，实业家不能离开统计而执业，政治家不能离开统计而施政。”统计思维是在获取数据、从数据中提取信息、论证结论可靠性等过程中表现出来的一种思维模式，对于人类提高认知起到巨大的作用。无论是解开自然奥秘的科学调查，或是考查早期匿名文学作品的作者、给出考古文



物的时间年表,或是解决法庭争端以及做出最佳决策等,统计思维都起到不可替代的重要作用。从获得知识的内在思维过程和决策行为本质来说,C.R.劳教授认为“统计学是人类探求真理必不可少的工具”。

统计学是一种由经验到理性的认识,是一种运用偶然发现规律的科学。它不只是一种方法或技术,还含有世界观的成分——看待世界上万千事物的一种方法,人们常讲某事从统计角度看如何,指的就是这个意思。统计思维的养成不但需要学习一些具体的指示,还要能够从发展的眼光,把这些指示连缀成一个有机的、清晰的图景,获得一种历史的厚重感(陈希孺,2002)。正如德国的斯勒兹曾讲道“统计是动态的历史,历史是静态的统计。”

从统计学的角度看,人们从经验或实验中所获取的知识是含有不确定性的,统计关注的是这种知识当中所含不确定性的度量问题,一旦能得到不确定性的量度,人们的知识就得到扩充,对世界的认知就朝前跨越,这个过程在人类知识积累的进程中不断重复。所以,C.R.劳教授总结道:“在终极的分析中,一切知识都是历史;在抽象的意义下,一切科学都是数学;在理性的基础上,所有的判断都源于统计学。”

三、透过统计思维理解科学事实的应有态度

统计方法作为一种实证主义的方法,目的是从偶然性中发现必然性,对真理做出探究。当今,在人类活动努力的一切范围内,统计学已经成为一种万能的、强有力的和不可缺少的研究工具(C.R.劳,2004)。统计作为探求真理的一种手段、工具,其效果如何还取决于运用的过程。

现实中源于多种原因,存在对统计有意或无意的误用、滥用,从而造成了一些统计陷阱和统计谎言。达莱尔·哈夫在《统计陷阱》一书中展示了各种利用统计行骗的技巧,如今经济和社会实证分析中对统计数据的“断章取义”和片面使用成为较普遍的现象

(朱玲,2002)。无论是对自然现象还是人文社科现象的分析研究,随便提出几个例子和片面寻找几个数据以佐证某种事物优于另一种事物的研究方式,不仅违背统计学研究方法中系统性原则,也与一般科学研究的方法论背道而驰。面对众多的统计谎言与陷阱,如何练就火眼金睛?达莱尔·哈夫在书中提出识破统计谎言的方法,即对统计资料提出五个问题:谁说的?他是如何知道的?遗漏了什么?是否有人偷换了概念?这个资料有意义吗?通过寻找这5个问题的答案,读者能够初步确定统计资料是否真实可信。识别统计谎言和陷阱,进而避免和防止统计欺骗的产生,增强统计探求真理的功能,还需要全方位的考虑和多方面的共同努力。

首先,数据采集及整理时要特别注意数据质量问题,避免操作失误和人为篡改,防止由错误的数据引出荒谬的结论。如果原始数据的质量有问题,统计分析只能是“垃圾进,垃圾出”。应用研究学者须对数据质量有一个清醒的判断,考虑到数据质量对实证结果可能产生的影响。数据处理必须通过严谨细密的数据录入、预处理、分析和检验工序构筑基础,加强过程的质量控制,否则即使拥有再高超的统计技巧,也难免会不自觉地制造出统计谎言。

其次,根据研究目的,结合数据的类型及特点,选择合适的统计方法。开展统计分析之前,必须认真考虑研究的主要目的是什么,希望解决什么问题,有哪些待检验的基本假设;分析过程中应针对资料类型和具体问题,从现有的方法库中选择合适的统计方法,如果现有方法不能满足要求,可以根据实际需要开展方法上的创新。

最后,不轻信自己的分析结果,多用常识和业务知识去检验。很多时候,统计分析方法都含有一些潜在的假设条件,实际分析过程中往往被忽略,故不要轻信统计分析结论,尤其是不能给自己的分析结论牵强地寻找理由!统计分析不能脱离背景知识和客观规律,如果得到与理论预设相反

的统计结果,应回头检查自己的假设条件,更深入地多方论证,而不能为了证明预先设想而摆弄统计游戏。

此外,统计结果的发布要全面、真实,而不能利用一些统计技巧隐藏真实的结果、有选择性地给用户以假象和诱导。

为达到对事物的全面、正确的科学认识,还应把以统计学为主的实证方法与规范方法结合起来使用。以规范分析为基础,才能保证统计等实证方法应用的有效性和适用性;实证分析的结果也只有进一步利用规范方法加以深入分析和研究,才能充分阐明数字背后科学含义和价值判断。□□

参考文献

- [1]陈希孺.数理统计简史.湖南教育出版社,2002年版.
- [2]程开明.统计:是真理,还是谎言.中国统计,2006(2).
- [3]C.R.劳.统计与真理——怎样运用偶然性(中文版).科学出版社,2004.
- [4]达莱尔·哈夫.统计陷阱(中文版).上海财经大学出版社,2002.
- [5]果壳译者.人人都该知道的20个科学事实,果壳网<http://www.guokr.com/article/437633/>,2013-11-28.
- [6]李金昌.统计思想研究.中国统计出版社,2009.
- [7]李竹渝、鲁万波.统计学与获取新知识.统计研究,2005(5).
- [8]Nature Editors. Number crunch——The correct use of statistics is not just good for science — it is essential. *Nature*, 2014, 506, (13): 131-132.
- [9]William J. Sutherland, David Spiegelhalter and Mark A. Burgman. Twenty tips for interpreting scientific claims[J]. *Nature*, 2013, 503, (21): 335-337.
- [10]朱玲.经验研究中的关键细节.经济研究,2002(11).

作者单位:浙江工商大学统计与数学学院