

心理学研究中假设检验 理论方法探析

焦璨 著

Jiao Can's Psychological Hypothesis Testing

中国社会科学出版社



深圳大学学术著作出版基金资助

Subsidized by Shenzhen University Foundation for the Production of Scholarly Monographs

心理学研究中假设检验 理论方法探析

焦璨 著

中国社会科学出版社

图书在版编目(CIP)数据

心理学研究中假设检验理论方法探析/焦璨著. —北京:中国社会科学出版社, 2014. 4

ISBN 978 - 7 - 5161 - 4256 - 1

I. ①心… II. ①焦… III. ①假设检验—应用—心理学研究方法—研究 IV. ①B841. 4

中国版本图书馆 CIP 数据核字(2014)第 091008 号

出版人 赵剑英
责任编辑 田 文
特约编辑 周 群
责任校对 王雪梅
责任印制 李 建

出 版 中国社会科学出版社
社 址 北京鼓楼西大街甲 158 号(邮编 100720)
网 址 <http://www.cssw.cn>
中文域名:中国社科网 010 - 64070619
发 行 部 010 - 84083685
门 市 部 010 - 84029450
经 销 新华书店及其他书店

印 刷 北京君升印刷有限公司
装 订 廊坊市广阳区广增装订厂
版 次 2014 年 4 月第 1 版
印 次 2014 年 4 月第 1 次印刷

开 本 710 × 1000 1/16
印 张 11.5
插 页 2
字 数 180 千字
定 价 35.00 元

凡购买中国社会科学出版社图书,如有质量问题请与本社联系调换

电话:010 - 64009791

版权所有 侵权必究

目 录

第一章 国内虚无假设检验基本使用情况	(1)
第二章 虚无假设检验的基础理论与发展沿革	(10)
第一节 虚无假设检验的相关概念	(10)
第二节 虚无假设检验的历史	(11)
第三节 虚无假设检验的发展	(16)
第三章 总体研究设计	(22)
第一节 研究设计	(22)
第二节 研究工具	(23)
第三节 Bootstrap 技术	(23)
第四节 与研究设计有关的概念	(24)
第四章 研究一:虚无假设检验新模式的构建	(30)
第一节 研究目的	(30)
第二节 统计功效的理论基础及其应用	(30)
第三节 效果量的理论基础及其应用	(36)
第四节 虚无假设检验新模式的构建及实例分析	(42)
第五章 研究二:样本容量、样本分布对单样本 t 检验的影响	(46)
第一节 研究目的	(46)
第二节 研究工具	(46)
第三节 抽样思路	(46)
第四节 检验指标和研究预期	(47)

第五节 研究过程及研究结果	(48)
第六节 抽样分布	(90)
第七节 研究结果与分析	(91)
 第六章 研究三:样本容量、样本分布对独立样本 t 检验的影响	(94)
第一节 研究目的	(94)
第二节 研究工具	(94)
第三节 抽样思路	(94)
第四节 检验指标和研究预期	(95)
第五节 研究过程及研究结果	(96)
第六节 研究结果与分析	(136)
 第七章 总结	(139)
第一节 主要结论	(140)
第二节 创新性	(143)
第三节 建议	(144)
 附录1 常见的效果量指标及标准	(146)
附录2 实例分析的统计功效分析表	(147)
附录3 研究二模拟程序	(148)
附录4 研究三模拟程序	(160)
附录5 自编程序验证	(169)
 参考文献	(172)
 后记	(178)

第一章

国内虚无假设检验基本使用情况

高等院校心理学专业自 1978 年恢复招生以来，特别是近 20 年来，心理学在中国得到长足的发展和广泛的应用。

随着中国社会经济的全面发展，国人对心理学的期许陡然增多，表现在：政府对心理学专业建设的投入增加，多个国家级、省市级心理学重点实验室、基地相继建立，核磁共振、眼动仪等各种高端设备投入使用，有关心理学的国家级、省部级等科研立项亦如雨后春笋，各类研究成果（以发表论文作为衡量标准）呈几何级数增长，可谓成果丰硕。这些成果大致可以分为认知心理学、神经心理学、教育心理学、学习心理学、社会心理学、人格心理学、心理学研究方法等方向。但遗憾的是：能解决社会重大问题和有重大影响的成果并不多见。例如，《心理学报》2008 年出版 12 期，发表研究论文共 140 篇，其中关于认知心理学方面的论文就有 66 篇，占全部发表文章总数的 47.14%，是成果最多的研究方向之一。大部分的研究集中在识字、阅读等方面。据发表论文所标识的项目资助统计，这 66 篇论文中有 90.91% 受到国家自然科学基金、国家社会科学基金、全国教育科学“十一五”规划重点课题、省部级科研立项等基金或项目的资助。这些都表明了认知心理学受到高度的关注，获得了巨大的学科建设投入，也表明了政府、社会对教育教学研究的高度重视。而一个巨大的反差是：社会、教育行业对这些成果的响应度不高，学校和一线教师对这些成果的认可程度也不高，很少应用于指导教师教学或学生学习。中小学生的识字能力、阅读能力并未因为这些研究成果的发表而提高，培养学生基本知识和技能的课时未见减少、学制未见缩短。学生学习压力大、“减负”问题依然是义务教育阶段的主要矛盾。从这一案例可知，心理学在社会发

展迅速的今天，离社会期许还有很大的距离，还未发挥应有的社会服务功效。

不能为社会实践所用的科学研究成果需要我们去考究其原因。科学的本质就在于可重复、可推广，而不囿于当前的研究被试。心理学作为一门应用学科，各领域、行业的认同以及研究成果的广泛推广，是其立命之本。为什么成果丰硕，实际应用却很少，这值得心理学界思考。出现论文大量发表，成果应用不多这一现象的原因有多种，包括研究变量确定、被试选取、实验设计、研究过程中额外变量的控制、研究方法的使用等方面。不同领域的心理学研究者都应反思，从多方面查找原因，审视各自存在的问题，担负应有的学术职责，以促进心理学科的发展。

心理统计方法作为心理学研究的重要工具之一。方法应用的正确与否对于研究成果的可靠性、可推广性很重要。虚无假设检验^①作为心理学研究中应用最多的重要统计方法，在中国心理学研究的应用情况如何？其应用是否正确？研究方法的应用是否影响研究成果的可靠性、科学性？这些都值得心理统计学的教学、研究工作者进行深入研究。本研究关注的就是心理统计分析方法的应用问题。

为了分析、研究中国心理学研究成果中统计分析方法应用的基本情况，总结各种心理统计方法的应用，寻找存在的问题及原因，并提出相应的解决方法及措施，本研究对中国最具影响力的两种心理学杂志：《心理学报》、《心理科学》（1998—2007）虚无假设检验的应用情况作元分析。

表1-1是《心理学报》、《心理科学》（1998—2007）虚无假设检验的使用总情况。从表1-1可知，两种杂志10年间发表的论文数有4115篇，使用虚无假设检验方法以佐证研究结论的有2752篇，占67%，其中，《心理学报》的使用率为84.8%，《心理科学》的使用率为61.2%。

^① 使用虚无假设检验（Null statistical hypothesis testing）主要区别于后文提到的Neyman-Pearson的假设检验（Hypothesis testing）。

表 1-1 《心理学报》和《心理科学》(1998—2007) 虚无假设检验使用总情况

	《心理学报》	《心理科学》	总数
总的文章数	989	3126	4115
入选文章数	839	1913	2752
百分比	84.8%	61.2%	67%

为了深入了解各种心理统计方法的应用情况,对各种方法的应用再作细分后统计,结果见表 1-2、表 1-3。表 1-2、表 1-3 分别是《心理学报》、《心理科学》(1998—2007) 发表的研究论文中各种虚无假设检验方法的基本使用情况。

表 1-2 《心理学报》(1998—2007) 虚无假设检验基本使用情况

年份	指标	总文章数	入选文章数	相关分析	t 检验	方差分析	协方差分析	回归分析	卡方检验	非参数检验	多元方差分析	因素分析	聚类分析	路径分析	判别分析	结构方程模型
2007	N	126	106	42	47	76	2	18	12	2	8	22	4	9	1	12
	t/a (%)			40	44	72	2	17	11	2	8	21	4	8	1	11
2006	N	111	98	32	23	63	1	15	14	5	8	28	0	2	0	15
	t/a (%)			33	23	64	1	15	14	5	8	29	0	2	0	15
2005	N	112	105	30	43	61	2	14	13	7	15	19	1	2	1	19
	t/a (%)			29	41	58	2	13	12	7	14	18	1	2	1	18
2004	N	103	85	16	36	62	1	8	14	4	9	15	3	3	0	6
	t/a (%)			19	42	73	1	9	16	5	11	18	4	4	0	7
2003	N	118	104	35	46	62	2	15	15	8	8	10	3	3	1	11
	t/a (%)			34	44	60	2	14	14	8	8	10	3	3	1	11
2002	N	98	85	24	37	49	0	11	17	2	13	8	3	0	0	8
	t/a (%)			28	44	58	0	13	20	2	15	9	4	0	0	9
2001	N	93	78	18	24	49	2	6	15	4	4	10	1	1	0	1
	t/a (%)			23	31	63	3	8	19	5	5	13	1	1	0	1
2000	N	80	72	16	28	48	0	6	13	0	9	5	3	0	0	1
	t/a (%)			22	39	67	0	8	18	0	13	7	4	0	0	1

4 心理学研究中假设检验理论方法探析

续表

年份	指标	总文章数	入选文章数	相关分析	<i>t</i> 检验	方差分析	协方差分析	回归分析	卡方检验	非参数检验	多元方差分析	因素分析	聚类分析	路径分析	判别分析	结构方程模型
1999	<i>N</i>	68	53	11	22	38	2	5	4	1	10	3	0	0	0	12
	<i>t/a</i> (%)			21	42	72	4	9	8	2	19	6	0	0	0	4
1998	<i>N</i>	80	53	10	34	33	3	7	4	2	7	5	1	1	0	0
	<i>t/a</i> (%)			19	64	62	6	13	8	4	13	9	2	2	0	0
总数	<i>N</i>	989	839	234	340	541	15	105	121	35	91	125	19	21	3	75

注：*t/a* (%) 表示使用某方法的文献占入选文献的比例；如 2007 年《心理学报》使用 *t* 检验的文献有 47 篇，占入选文章数（106 篇）的 44%。由于同一篇文章使用的虚无假设方法可能不止一种，所以有的文章进行了多次编码，*t/a* (%) 的总和故不为 100%。

表 1-3 《心理科学》（1998—2007）虚无假设检验基本使用情况

年份	指标	总文章数	入选文章数	相关分析	<i>t</i> 检验	方差分析	协方差分析	回归分析	卡方检验	非参数检验	多元方差分析	因素分析	聚类分析	路径分析	判别分析	结构方程模型
2007	<i>N</i>	405	277	68	83	128	2	47	45	3	19	52	4	3	2	17
	<i>t/a</i> (%)			25	30	46	1	17	16	1	7	19	1	1	1	6
2006	<i>N</i>	401	275	113	120	158	4	46	28	6	19	70	4	3	1	12
	<i>t/a</i> (%)			41	44	57	1	17	10	2	7	25	1	1	0.4	4
2005	<i>N</i>	416	255	107	114	150	5	46	33	9	19	44	9	9	1	9
	<i>t/a</i> (%)			42	45	59	2	18	13	4	7	17	4	4	0.4	4
2004	<i>N</i>	440	260	83	104	128	5	34	42	0	25	49	4	2	1	15
	<i>t/a</i> (%)			32	40	49	2	13	16	0	10	19	2	1	0.4	6
2003	<i>N</i>	378	221	87	100	131	1	32	15	4	16	37	8	4	3	5
	<i>t/a</i> (%)			39	45	59	0	14	7	2	7	17	4	2	1	2
2002	<i>N</i>	262	157	42	71	76	2	19	18	3	8	25	0	2	0	1
	<i>t/a</i> (%)			27	45	48	1	12	11	2	5	16	0	1	0	1

续表

年份	指标	总文章数	入选文章数	相关分析	t 检验	方差分析	协方差分析	回归分析	卡方检验	非参数检验	多元方差分析	因素分析	聚类分析	路径分析	判别分析	结构方程模型
2001	N	266	144	47	56	80	3	16	18	1	13	24	4	2	2	4
	t/a (%)			33	39	56	2	11	13	1	9	17	3	1	1	3
2000	N	217	129	43	57	70	1	12	19	0	6	20	6	1	1	1
	t/a (%)			33	44	54	1	9	15	0	5	16	5	1	1	1
1999	N	176	94	15	45	56	2	6	10	1	13	10	0	2	0	0
	t/a (%)			16	48	60	2	6	11	1	14	11	0	2	0	0
1998	N	165	101	37	44	50	1	10	12	0	10	13	2	3	2	0
	t/a (%)			37	44	50	1	10	12	0	10	13	2	3	2	0
总数	N	3126	1913	642	794	1027	26	268	240	27	148	344	41	31	13	64

注：t/a (%) 表示使用某方法的文献占入选文献的比例；如 2007 年《心理科学》使用 t 检验的文献有 83 篇，占入选文章数（277 篇）的 30%。由于同一篇文章使用的虚无假设方法可能不止一种，所以有的文章进行了多次编码，故 t/a (%) 的总和不为 100%。

从表 1-2、表 1-3 可以看出，虚无假设检验方法中使用最多的是方差分析；t 检验次之；使用最少的是协方差分析和判别分析。可见，和其他国家的心理学研究一样，虚无假设检验在中国心理学研究中得到了广泛应用，已成为中国心理学研究的最重要的统计方法之一。

如前所述，虚无假设检验在心理学研究中广为应用，其显著性结论成为研究论文得以接受、发表的重要标准之一（注：唯有显著性结论的论文才能得到发表，这已成为心理学科的行业标准。可以预见：得到显著性结论未能发表和未得到显著性结论的研究为数更多，在所有的心理学研究中占据了很大的比例，这类研究常被束之高阁，产生强烈的“文件柜”效应）。应用了量化研究方法，并获得了“显著性差异”的结论，如此丰硕的心理学研究成果未得到广泛应用，从方法学的角度，究其原因，有如下几点：

第一，心理学研究对象的特殊性。心理学的研究对象是人，是人的各

种心理特质、心理现象，具有间接测量性、不可重复性。由于人具有学习能力及成长性，所以相对于其他自然科学实验，心理学研究的重复性、被试的选择等难度更大，对这种特殊的研究对象，统计方法的使用亦应有其特殊性。

第二，对心理学研究目的认识不足。心理研究工作者在研究中考虑了实验因素的可控性、取样的方便性。但研究目的仅停留在实验组和控制组之间、各因素及因素各水平之间是否存在差异，即差异的有无，获得统计上的显著性差异则止。未深刻意识到差异的大小所提供的信息才是真正的研究目的^[1]，导致研究成果与实际应用之间严重脱节。

第三，统计分析方法应用中存在问题。尽管大部分研究者已经意识到统计方法的重要性，认可统计分析方法在心理学研究中的地位和作用，并在研究中大量使用虚无假设检验。但是学术刊物对心理统计学、心理测量学等基础方法学研究领域并不重视，未强调统计方法应用的科学性要求，加上“傻瓜式”统计软件的鼠标式使用，使很多心理学研究者对心理统计学、心理测量学原理不求甚解，对各类虚无假设检验方法的使用前提视而不见，将统计分析方法直接引入各种研究过程。

第四，抽样缺乏科学性。抽样必定伴随着误差，误差有随机误差和系统误差之分。心理学研究者将样本异质性视为随机因素，将样本异质性产生的误差视为随机误差。实际上样本异质性产生的误差应该包含随机（异质性）误差和系统（异质性）误差。其中，随机异质性是指被试之间能力、心理特质的异质性，包括被试的天赋水平和努力程度等影响因素。系统异质性是指不同被试群体之间的异质性，包括被试所处的地区、民族、家庭背景、学校背景以及教育经历等影响因素。随机异质性和系统异质性共同影响样本异质性，所以抽样时应该同时考虑二者。目前，由于人力、物力、财力等因素以及取样的困难，心理学研究常进行方便抽样，采用学生样本。发表的心理学研究论文常采用诸如“在某高校随机抽取学生××名”、“在某地区随机抽取被试××名”等被试选取方案，用以研究大学生或者其他群体的某一心理现象。我们以实际案例予以说明。

《心理学报》2008年第1期共发表14篇论文。在研究结果来自实

测数据的 11 篇论文中, 6 篇论文的被试从某一学校抽取, 1 篇论文的被试从某两所学校抽取。忽视系统误差的抽样方式就占该期有抽样研究的论文的 64%。我们不禁要问: 这些学校的被试样本能代表被试总体吗?

这种抽样方法所抽取的样本存在的最大问题是无法获知其总体分布情况。而样本分布常表现为: 与假设的偏度和峰度均为零的正态总体相比较, 样本的偏度和峰度均不为零, 出现尖峰和重尾, 样本分布偏离正态分布。这种抽样方式不仅会产生随机异质性误差, 也会产生系统异质性误差, 而虚无假设检验因为方法应用的条件所限, 仅能分离随机误差、随机异质性, 不能分离系统误差、系统异质性。若随机异质性误差、系统异质性误差在一个实验中同时出现, 对虚无假设检验的结论产生多大影响, 在多大程度上会破坏研究结论的可靠性、可重复性?

为助于对这一问题有更为深刻的认识, 本研究对《心理学报》、《心理科学》(1998—2007) 发表论文中使用最多的两类虚无假设检验方法: t 检验、方差分析在应用中是否做前提假设检验进行统计, 结果见表 1-4。

表 1-4 《心理学报》、《心理科学》(1998—2007) t 检验、方差分析的前提检验情况

			N 检验/N 方法									
			2007	2006	2005	2004	2003	2002	2001	2000	1999	1998
《心理学报》	t 检验	总体正态分布					2/27					
		方差齐性	1/28		1/27		3/27	2/28				
	方差分析	总体正态分布	1/76	1/63								
		方差齐性	5/76	5/63						1/48		
《心理科学》	t 检验	总体正态分布			1/81	1/69						
		方差齐性	1/66	1/86	1/81							
	方差分析	总体正态分布		1/158	1/150							
		方差齐性	1/128	2/158	3/150	3/128	2/131		1/80	2/70		1/50

注: 以 N 检验/ N 方法作为指标, 即在使用方法的总次数中进行了前提检验的比率。

从表 1-4 中可知,严格按照统计分析要求进行前提检验的研究或实验很少,仅占 2.95%。忽略前提假设检验这一步骤会产生什么样的结果,不满足分布条件的样本数据会对研究结论产生怎样的影响?这一问题值得心理学研究者,特别是心理统计研究者深入研究分析,寻求答案,这亦是本研究的目的。

未正确理解、使用虚无假设检验,不管是对于中国心理学研究者,还是对于其他国家的心理学研究者而言都普遍存在^[2]。因此,自 20 世纪 70 年代开始,作为统计学的重要方法之一,虚无假设检验就遭到了心理统计学、数理统计学等领域研究者的很多批评^[3,4,5,6,7,8,9,10]。对样本容量的依赖性以及无法估计研究结论的可重复性、可推广性是研究者对虚无假设检验提出的最为主要、最为尖锐的批评。尽管心理统计研究者对这些批评做出了回应和反思,作为充实,对统计功效、效果量方面亦开展了研究,并取得了一定的成果,这些成果主要集中在:强调统计功效、效果量的重要性和必要性^[11,12,13,14,15],提出各类效果量指标^[16,17,18,19,20];对各杂志已发表论文的统计功效、效果量进行调查和元分析^[21,22,23,24,25];对各种效果量指标进行比较^[24];探讨样本容量对统计功效、效果量的影响等^[26,27],美国心理学协会(American Psychological Association, APA)也在发表论文审稿要求上作出了相应调整^[28,29]。但到目前为止,还没有完全解决假设检验模式构建的问题。

因此,本研究拟在前人研究的基础上,在虚无假设检验的理论、方法上作进一步深入的研究,并采用计算机数据模拟技术,达到以下研究目的:

第一,拟构建心理统计假设检验的新模式。以重视心理研究结论的可重复性作为目标,构建新的包括统计功效、显著性检验和效果量共存互证的虚无假设检验模式。从而提出一个既可以承载差异性有无的信息,也可以提供差异大小、结论可重复性信息,并能用以横向、纵向比较不同研究之间的虚无假设检验模式。

第二,在虚无假设检验的新模式下,采用计算机数据模拟的方法,选择单样本 t 检验、独立样本 t 检验这两种心理学研究使用广泛的统计模型进行深入研究。在已有的研究成果,即样本容量对虚无假设检验的显著性

结论、统计功效产生影响，效果量不受样本容量的影响的基础上，重点研究样本偏度、峰度作为样本分布特征对显著性结论和效果量的影响。这一问题的研究至今未见相关文献，仅 Finch 等人在“Colloquium on Effect Size: the Roles of Editors, Textbook, Authors, and the Publication Manual”一文中提出了对样本变异性会影响效果量这一问题的猜想和质疑^[30]。样本容量、偏度和峰度会在多大程度上对虚无假设检验的显著性结论和效果量产生影响，当样本容量、样本分布的偏度及峰度达到多大对二者产生的影响会趋于稳定，这是心理统计理论研究的重要的基础理论，亦是本研究的主要目的。

第二章

虚无假设检验的基础理论与发展沿革

了解虚无假设检验的概念、基础理论及发展沿革，有助于了解本研究的理论依据及构思。

第一节 虚无假设检验的相关概念

一 两类错误：Ⅰ型错误和Ⅱ型错误

心理学研究中使用虚无假设检验可能产生如下四种结果^[31]。

表 2 - 1 虚无假设检验中的两类错误

	未拒绝 H_0	拒绝 H_0
H_0 为真	$1 - \alpha$	α
H_0 不真	β	$1 - \beta$

虚无假设 H_0 为真而被拒绝以及虚无假设^① H_0 不为真而未能拒绝这两种情况下得到的是错误的结果。前者称“弃真”（Ⅰ型）错误，记为 α ，表示为： $\alpha = P\{\xi \in \omega \mid H_0\}$ （ ω 为 H_0 的拒绝域）；后者称“取伪”（Ⅱ型）错误，记为 β ，可表示为： $\beta = P\{\xi \notin \omega \mid H_1\}$ 。 H_0 为真而未被拒绝以及 H_0 不为真而被拒绝，这两种情况得到的是正确的结果。

^① 使用虚无假设（Null hypothesis）而不使用零假设这一传统术语，主要原因如下：零假设容易让读者认为 H_0 即意味着零差异、零效应；实际上 H_0 仅意味着当前实验对于证据的收集是无效的。

二 统计功效 (statistical power)

统计功效是推断统计中的重要概念,即备择假设 H_1 为真时,拒绝虚无假设 H_0 的概率^[11]。统计功效与犯Ⅱ型错误的概率 β 相联系,是 β 的余数,记为 $1 - \beta$ 。统计功效由下式定义:

$$1 - \beta = 1 - P\{\xi \in \omega \mid H_1\} = P\{\xi \in \omega \mid H_1\} \quad (2-1)$$

三 效果量 (effect size)

效果量是指虚无假设 H_0 不为真的程度,实际上就是备择假设 H_1 偏离虚无假设 H_0 程度的一种指数^[11]。效果量是统计功效诸多影响因素中最为重要的一个,常作为统计功效大小的直接量尺。

第二节 虚无假设检验的历史

1879年,冯特在德国莱比锡大学成立了世界上第一个心理学实验室,从此,心理学脱离哲学的襁褓终为一门独立的科学。20世纪初,随着 Fetcher 心理物理学的开创, Ebbinghaus 记忆遗忘曲线的发现,第一个心理测验——比纳西蒙智力量表的编制,以及统计学方法的引入,心理学得到发展,心理学量化研究也得到了长足的进步。目前心理学已成为一门定性研究和定量研究相结合,科学主义和人文主义相交融的边缘科学、综合科学。心理统计学、测量学以及心理实验设计在心理学不断科学化的发展过程中发挥着重要作用。心理统计学已成为各大院校心理学专业的必修课程。

心理统计学一般包括描述统计和推断统计。前者主要是研究、简缩和描述数据;后者主要是研究如何利用数据进行决策的方法,是进行心理学实验或对心理学研究作出预测、规划的有力工具。虚无假设检验 (null statistical hypothesis testing) 是推断统计的主体。心理学量化研究中,运用虚无假设检验佐证心理学研究成果的历史悠久,心理学的诸多重要研究都依靠虚无假设检验来作出判断。目前,虚无假设检验已是心理学研究,尤其是心理实验研究、心理测量、问卷调查等的重要研究工具和方法,也成为

了研究论文接受、发表的重要标准之一，在心理学不断科学化的过程中发挥着重要作用。

虚无假设检验是推断统计的重要组成部分。作为一种正规的统计方法，虚无假设检验是在同一数据基础上对两个或者两个以上假设统计模型的检验，是在“虚无假设（null hypothesis） H_0 为真”的假设前提下计算以往观测到的结果所发生的概率，其理论依据是“小概率事件在一次试验中几乎不可能发生”的原理。虚无假设检验为我们提供了拒绝某一假设的工具^[32]。

现有的一些数理统计学以及心理统计学、教育统计学等应用统计学教材对于虚无假设检验的介绍、讨论，使读者产生错觉，认为虚无假设检验是一套固有的理论，将之视为统一的、丝毫没有争议的理论方法。实则不然，目前我们广泛使用的虚无假设检验是种不完备的结合体，是 20 世纪二三十年代 Fisher 的显著性检验（significance testing）和 Neyman - Pearson 的假设检验（hypothesis testing）的结合体^[33]。这是现代统计学发展史上影响颇为重大且又相互矛盾的两大学派。尊为“现代统计学之父”的 Fisher 发展了从测量数据中产生显著性水平的方法，而 Neyman 和 Pearson 则提出了一套严格地拒绝某一事先确定的假设的决策程序。

一 Fisher 的显著性检验（significance testing）模式

Fisher 的显著性检验模式一般包括以下几个步骤^[34]：

- （1）确定假设 H ；
- （2）决定合适的检验统计量及其在 H 为真的前提下的分布；
- （3）从测量数据中计算检验统计量 T ；
- （4）使用 H 为真的前提下 T 的分布，确定与之相对应的显著性水平 P ；
- （5）如果获得的显著性水平非常小，则拒绝 H 。值得注意的是：这是显著性检验可以得到的唯一的结论。

在 Fisher 的理论框架中，仅设置了一个关于检验统计量 T 的已知分布的假设 H 。若该检验统计量及其条件期望值 $E(T|H)$ 相差甚远，假设则不大可能偶然发生。