

---

# P值的前世今生

作者：writer 来源：本站

本文原地址：<https://www.iikx.com/news/statistics/271.html>

**本文仅供学习交流之用，版权归原作者所有，请勿用于商业用途！**

p值可能是现代科学论述中使用最广泛的检验统计值。它经常被新研究生和读者错误地用来解读已知研究的全部实验数据。最初它是假设检验决策的有用工具，后来成为一验即灵的检验方法，以确定结果是否显著，甚至很不幸地沿用在文章是否发表的决策上。

因此，我们想对p值提供一个简短明了的描述，以及使用和解释它应有的正确方式。这篇文章将为作者提供p值作为当前科学领域的工具的最新解释。请注意，如果您的手稿中存在对p值的明显误解，那么您被审稿人立即拒稿也是合情合理的。

p值最初被计算为检验统计量，用来描述一组基于零假设的给定数据。皮埃尔·西蒙·拉普拉斯(Pierre-Simon Laplace)——也是对表面张力的用数学语言描述的提出者——原本在试图计算p值用来将性别分布进行归类为real。因此，这个概念起源于p值可以检测差异是否是真实存在，或者只是出于巧合。p值的效用，在于建立起一个基于样本数据结果，以判断拒绝或接受假设的通用的标准化决策过程。根据罗纳德·费舍尔(Ronald Fisher)的建议，将零假设的拒绝阈值设置为

那么，考虑到p值的效用，究竟是在计算什么？

p值是对数据的描述;这不是对这个假设的描述。该值表示，在假设零假设为真实情况下，试验结果为极端数据的概率。这是决定是拒绝还是接受零假设的有用工具。科学家们已经达成了拒绝零假设的阈值的共识。这直接反映了错误地拒绝零假设(第一类错误)或错误地接受零假设(第二类型错误)的概率。因此，它是判断明显差异为real的可能性的直观指标。

不幸的是，这种直观性经常导致p值的普遍性滥用，最近的研究进展，如再现性危机，已经改变了对p值的使用和报告的态度。了解这些变化对于文章的成功发表至关重要。

最近，由于p值的争议不断，美国统计局觉得发表一个关于使用p值的声明迫在眉睫。

以下是摘自美国统计局对p值的声明[1]：

“ P values do not measure the probability that the studied hypothesis is true, or the probability that the data were produced by random chance alone. ”

“ p值不测量研究假设是真实的概率，或者样本源自随机数据的概率。 ”

“ Scientific conclusions and business or policy decisions should not be based only on whether a p value

---

passes a specific threshold. ”

“ 科学结论、商业决策或者政策制定不应仅仅基于p值是否通过特定的阈值。 ”

“ A p value, or statistical significance, does not measure the size of an effect or the importance of a result. ”

“ p值，或统计学显著性，不测量效应的大小或结果的重要性。 ”

这三个判断对现代如何使用p值至关重要。现在我们将根据上述信息为您的手稿中使用p值提供一些指导。

1、p值越来越多地被报告为精确值(比如， $p=0.012$ )而不是阈值(p值)

2、p值不能单独报告。鉴于p值不提供关于结果的效应量(effect size)、可概括性(generalizability)或重要性的证据，审稿人期望看到相应的统计证据。这可以包括报告效应量，置信区间(confidence intervals)和标准误差。同样地，由于上述原因，使用诸如highly significant “ 高度显著 ” 的短语是不合适的。

3、不要试图规避统计学显著性。基于对p值被广泛滥用于确定一项研究是否成功或失败的共识，审稿人对试图篡改其p值的作者是不予原谅的。这种现象通常被称为p-hacking “ p值篡改 ”。下面举两个例子——大家会发现时不时在一些地方看到：1)先说实验数据发现了一个差异，随后再接着说but this difference did not reach statistical significance.( “ 但是这个差异没有达到统计学意义。 ” );2)使用诸如 “ marginally significant ” 边缘重要性这样的短语。这两种写法都是错误而不可取的。

4、假设陈述应尽可能具体。p值基于归谬法(reduction ad absurdum logic)。因为零假设被认为是不可能的，备择假设或实验假设则会被接受。维护这一逻辑结构至关重要;因为只存在两个解释，只有备择假设和零假设是可能的。因此，零假设通常是无效的假设。确保您的备择假设是一个合适的对立陈述。

更多 统计方法 请访问 <https://www.iikx.com/news/statistics/>

本文版权归原作者所有，请勿用于商业用途，[爱科学iikx.com](http://www.iikx.com)转发