

【编委之窗】

假设检验的一个常见误区

韩兆洲, 魏章进

(暨南大学 经济学院统计学系, 广东 广州 510632)

摘 要: 文章通过对假设检验的一个常见错误进行了理论分析, 指出假设检验的方法只能在一定情况下否定原假设而不能肯定原假设, 并提出了设定原假设和备择假设的正确而简明的方法。

关键词: 假设检验; 原假设; 备择假设

中图分类号: O212.1 **文献标识码:** A **文章编号:** 1007-3116(2005)01-0009-03

一、问题的提出

在统计学教学实践中, 许多人对于下面的假设检验问题常常会得出一个令人困惑的结论。问题如下: 某厂欲从外地进口一批家用电器, 该生产厂家推销员声称其产品有 95% 以上的合格率, 现从其中抽取 200 件样品, 发现共有 8 件不合格品, 问是否应该相信该推销员的说法(显著性水平 $\alpha = 0.05$)。

对此问题的解法常见的有以下两种:

解法 1 设原假设 $H_0: P \leq 0.95$

备择假设 $H_1: P > 0.95$

大样本情况下用正态分布近似^[1], 在 $\alpha = 0.05$ 时, $u_{0.95} = 1.645$, 拒绝域应为 $\{u \geq 1.645\}$, 现由抽样结果求得:

$$\begin{aligned} u &= \frac{\sum x - np}{\sqrt{np(1-p)}} \\ &= \frac{192 - 200 \times 0.95}{\sqrt{200 \times 0.95 \times 0.05}} = 0.65 < 1.645 \end{aligned}$$

观测值未落入拒绝域, 则结论为不能拒绝原假设, 即不能断定该生产厂家的产品的合格率在 95% 以上, 从而不能断定推销员的说法的正确性。

解法 2 设原假设 $H_0: P \geq 0.95$

备择假设 $H_1: P < 0.95$

大样本情况下用正态分布近似, 在 $\alpha = 0.05$ 时, $u_{0.95} = 1.645$, 拒绝域应为 $\{u \leq -1.645\}$, 现由

抽样结果求得:

$$\begin{aligned} u &= \frac{\sum x - np}{\sqrt{np(1-p)}} \\ &= \frac{192 - 200 \times 0.95}{\sqrt{200 \times 0.95 \times 0.05}} = 0.65 \end{aligned}$$

观测值也未落入拒绝域, 则结论为接受原假设, 即其产品的合格率在 95% 以上, 从而相信该推销员的说法。

两种方法出现了“相反”的结论。这是一个令初学者十分头痛的问题。那么, 为什么会出现所谓“相反”的结论? 什么样的假设检验才是正确的假设? 以及如何在类似情况下进行假设检验呢? 本文将对这一问题进行探讨。

二、出问题的原因

为了搞清楚问题的原因所在, 有必要了解假设检验的真正作用, 即假设检验到底能检验什么? 题中检验的置信水平为 95%, 建立的统计量 u 在原假设成立的情况下, 根据概率论的原理 u 处于接受区域的概率为 95%。由于抽样的随机性, 随机样本 u 有可能落入拒绝域, 而在原假设成立的情况下, 它落入的概率仅为 5%, 而在抽样中若它真的出现了, 将之归结为小概率事件。笔者认为: 小概率事件是不可能发生的, 它居然发生了, 这说明了什么呢? 只能说明原假设并不成立, 备择假设成立。这种结论的可信度

收稿日期: 2004-10-08

作者简介: 韩兆洲(1955-), 男, 江苏省苏州市人, 教授, 博士, 博士生导师, 研究方向: 经济预测与决策;

魏章进(1966-), 男, 湖北省洪湖人, 硕士生, 研究方向: 经济预测与决策。

有多高呢?有 95% 的把握, 为什么这样说呢?因为原假设成立, 则 95% 的情况下样本都不会落入拒绝域。换言之也就是不会拒绝原假设, 这样在原假设成立而拒绝原假设的可能性只有 5%, 即拒真的概率只有 5%; 反之, 拒绝原假设有 95% 的可信度, 于是可以认为在假设检验中一旦拒绝原假设, 则可能有相当大的把握。亦即假设检验在否定原假设时是令人放心的, 即假设检验能检验备择假设的真实性。

若样本 u 落入了接受域, 此时假设检验将说明什么呢?它的真正含义值得探讨。样本落入了接受域是否意味着原假设成立呢?其实并非如此, 因为由于样本的随机性, 原假设不成立时样本也有可能落入接受域, 即原假设不成立而接受原假设, 这种错误即取伪的错误。而取伪的概率在假设检验中并不能具体给出。那么样本一旦落入接受域又意味什么呢?如果样本落入拒绝域, 则可以声称备择假设 95% 的情况成立; 若没有落入拒绝域, 而落入接受域只是意味着还没有 95% (本题中显著性水平 5%) 的把握肯定备择假设否定原假设, 即落入接受域并不意味原假设成立, 只是意味着不能肯定地否定原假设。此时接受原假设就会冒取伪的风险, 假设检验并不能验证原假设的真实性。

综上所述, 若样本落入拒绝域, 则可以肯定原假设不成立; 若样本落入接受域, 则表明还不能以一定的置信水平否定原假设成立。如果不能否定原假设时, 且并不严格地接受原假设, 这种不严格的说法在一些教科书中常见, 是因为有一个前提, 即把已经公认的广为接受的问题设为原假设, 就不能有充分把握否定它, 那就接受它, 这就是所说的保护原假设的缘由所在。明白了这些严格的结论, 就可以解释上述问题了。解法 1 的结论: 没有 95% 的把握断定产品的合格率高于 95%。解法 2 的结论: 没有 95% 的把握断定产品的合格率低于 95%。所以两个结论并不矛盾。所谓相反的结论是由于不严格的说法所导致的。由于许多教科书中不太严格的说法导致了初学者误以为此时原假设成立, 但实际上正确的说法是不能否定原假设。而是否可以接受还需要进一步检验, 不能否定隐含着需要进一步检验的意思。

三、问题是非的辨析

正确的假设首先应该是检验的结论能达到检验之目的。从假设检验的目的来看, 解法 1 显然达到了我们的目的, 并明确地说明: 没有 95% 的把握认定该厂的家电产品的合格率高于 95%。而解法 2 得到

的结论并没有达到我们的目的, 因为不能断定合格率低于 95%, 并不等于合格率高于 95%; 再者不能否定原假设则并不意味着原假设成立, 如解法 2 中 u 落入接受域, 如果说不能否定之, 假如接受它则有可能取伪, 即犯第二类错误。由于抽样的随机性, 在假设不真的情况下它也可能落入接受域, 在检验中一旦接受原假设, 我们就必须考虑取伪概率。因此看一下解法 1 和解法 2 的取伪概率(样本合格率为 p_0 , 总体合格率为 P):

解法 1 的取伪概率为:

$$P = P\left\{\frac{\sum x - np_0}{\sqrt{np_0(1-p_0)}} < u_{0.95} = 1.645/H_0 \text{ 非真}\right\}$$

$$\approx P\left\{\frac{\sum x - np}{\sqrt{np(1-p)}} < u_{0.95} - \frac{n(p-p_0)}{\sqrt{np(1-P)}}\right\}$$

$$= \Phi\{u_{0.95} - (p-p_0)/\sqrt{p(1-P)/n}\}$$

P 若远远大于 p_0 则取伪概率就会变小, 实际上本例中两者接近, 取伪概率很大。

解法 2 中的取伪概率为:

$$P = P\left\{\frac{\sum x - np_0}{\sqrt{np_0(1-p_0)}} > -u_{0.95} = 1.645/H_0 \text{ 非真}\right\}$$

$$\approx P\left\{\frac{\sum x - np}{\sqrt{np(1-p)}} > u_{0.95} + \frac{n(p-p_0)}{\sqrt{np(1-P)}}\right\}$$

$$= \Phi\{u_{0.95} + (p-p_0)/\sqrt{p(1-P)/n}\} > 0.95$$

可见解法 2 的取伪概率更大, 大于 95%。在这样的情况下接受原假设成立的结论显然是十分荒谬的。

如前已经看到在解法 1 中 P 若远远大于 p_0 , 则取伪概率就会变小, 而实际上两者接近取伪概率也很大。但为什么解法 1 正确呢?因为解法 1 中的检验并没有试图去承认原假设, 它的结论是不能否定原假设, 不需要考虑取伪概率。因为拒真的错误已被显著性水平限定, 所以得到否定原假设从而接受备择假设的结论是可靠的; 而解法 2 要达到检验的目的, 则需要接受原假设, 而这时就很有可能犯下取伪的错误。本题中取伪概率就很大, 而且大于 95%。正是这一点才是解法 1 为什么正确, 而解法 2 为什么错误的關鍵。

四、解决问题的原则

综上所述, 假设的原则应该是把希望验证得到的结论放在备择假设的位置上。利用这一原则, 对解决这一类问题就非常有效, 它把处理参数假设检验的问题推广到更为一般的情况。

例如问题: 微波炉在炉门关闭时的辐射量是一

个重要的质量指标。某厂该指标服从正态分布 $N(\mu, \sigma^2)$, 长期以来 $\sigma = 0.1$, 且均值都符合要求不超过 0.12。为检查近期产品的质量, 抽查了 25 台, 得到炉门关闭时辐射量的均值为 0.123。试问在 $\alpha = 0.05$ 的水平上该厂炉门关闭时辐射量是否升高了?

在此题中, 如何来设立原假设呢? 由于题目中所问辐射量是否升高, 那么需要验证的结论即是 μ 是否大于 0.12, 故应将 $\mu > 0.12$ 作为备择假设。

故解为: 首先建立假设 $H_0: \mu \leq 0.12, H_1: \mu >$

0.12。在 $\alpha = 0.05$ 时, $u_{0.95} = 1.645$, 拒绝域应为 $\{u \geq 1.645\}$ 。现由观测值求得 $u = 0.015 < 1.645$, 因而在 $\alpha = 0.05$ 的水平下, 不能拒绝 H_0 , 即认为当前生产的微波炉关门时的辐射量无明显的提高。

在统计学的假设检验的教学过程中, 这个问题的重要性无论怎么强调都不过份。把上述的问题阐述清楚, 有助于学生对假设检验的透彻理解, 并对后续的理论学习大有裨益。

参考文献

[1] 茆诗松, 周继蓁. 概率论与数理统计[M]. 北京: 中国统计出版社, 1999.

(责任编辑: 郭诗梦)

A Frequent Mistaken Concept in Hypothesis Test

HAN Zhao-zhou, WEI Zhang-jin

(Dept. of Statistics, College of Economics, Jinan University, Guangzhou 510632, Guangdong)

Abstract: Based on the theoretic analysis of a familiar error of hypothesis testing, this paper explains an only situation that we can deny the null hypothesis, and puts forward a accurate way to make the null hypothesis and the alternative hypothesis.

Key words: Hypothesis testing; Null hypothesis; Alternative hypothesis.

《统计与信息论坛》2005 年元旦茶话会 在西安统计研究院举行

由《统计与信息论坛》编辑部组织的 2005 年元旦茶话会于 2004 年 12 月 29 日在西安统计研究院会议交流中心举行。来自西安财经学院、西安交通大学等在西安的编委、审稿专家 30 多人参加了此次会议, 西安财经学院院长兼西安统计研究院院务会副主席杨学义教授以及西安财经学院经济学院、信息学院、管理学院、学科建设办公室的负责同志也应邀到会。

茶话会由主编刘延年教授主持。首先由杜一哲同志介绍了《统计与信息论坛》一年来的办刊情况, 之后, 各位专家对《统计与信息论坛》一年来所取得的进步予以充分肯定, 同时也对如何进一步办好杂志提出了许多宝贵的建议和意见。在漫谈中, 大家亦就统计学的发展趋势及其科学研究方面的问题进行了探讨, 就提高审稿速度和审稿质量交换了看法和经验。会上, 王振龙副主编还就《统计与信息论坛》如何迈上一个更新的台阶介绍了进展情况。

茶话会洋溢着轻松温馨的气氛, 大家畅所欲言, 对《统计与信息论坛》的未来发展充满了信心。

(一哲)