

假设检验的原理及其应用

王志福, 潘旭, 金姝, 田丰, 王艳

(渤海大学 数理学院, 辽宁 锦州 121013)

摘要: 假设检验是数理统计中的重要内容. 文章介绍了假设检验的原理和结果, 简单阐述了原假设和备择假设的设置原则, 并应用单向等级方差分析和双向等级方差分析两种假设检验方法解决非参数检验问题, 简化了计算过程.

关键词: 假设检验; 非参数检验; 单向等级方差分析; 双向等级方差分析

中图分类号: O212.1 **文献标志码:** A **文章编号:** 1673-0569(2013)02-0101-05

0 引言

数理统计学的一个基本问题就是从所要研究的对象全体中抽取一部分进行观测或实验, 根据所取得的有限信息对整体进行推断的问题, 每个推断必须伴随一定的概率以表明推断的可靠程度, 这种伴随有一定概率的推断称为统计推断, 统计推断中的一个主要方面就是统计假设检验. 假设检验包含参数检验和非参数检验两大类, 文章中就以非参数检验为例, 应用单向等级方差分析和双向等级方差分析两种假设检验方法解决实际问题, 从而简化了数据处理过程.

1 假设检验的基本概念

统计假设是任意一个有关未知分布的假设, 它是一种关于母体分布状态的陈述, 也可简称为假设. 如果一个统计假设完全确定母体的分布, 那么称这种假设为简单统计假设或简单假设, 否则就称为复合统计假设或复合假设. 在假设检验中, 一般提出两种假设: 原假设 H_0 和备择假设 H_1 , 当我们对所构造的法则作出判断时, 拒绝原假设 H_0 还是拒绝备择假设 H_1 , 这种法则就称为 H_0 对 H_1 的一个检验法则⁽¹⁾, 有时就简称为一个检验法则, 或一个检验, 检验法则本质上就是把子样空间 Ω 划分为两个互不相交的子集 A 和 A^* . 参数假设是仅仅涉及到母体分布的未知参数的统计假设, 而非参数假设则只对未知分布函数的类型或者它的某些特征提出某种假设. 在数理统计学中把不依赖于分布的统计方法统称为非参数统计方法, 这里所讨论的问题就是非参数假设检验问题.

2 假设检验的基本原理

假设检验的基本思想是依据过去的信息对不能肯定(确定)的问题做出一个肯定(或否定)的回答. 假设检验的基本原理: 在给定备择假设 H_1 下对原假设 H_0 作出判断, 若拒绝原假设 H_0 , 那就意味着接受备择假设 H_1 , 否则就接受原假设 H_0 . 这就是依据小概率事件原理的反证法⁽²⁾.

收稿日期: 2013-03-14.

基金项目: 辽宁省自然科学基金资助项目(No: 20101001); 辽宁省教育厅自然科学基金资助项目(No: 2008L038).

作者简介: 王志福(1960-), 男, 教授, 从事应用概率与统计方面研究.

通讯作者: panxuvictoria@163.com.

在假设检验中,我们构造某一个事件使它在原假设 H_0 为真的条件下发生的概率很小,现在在一次试验或观察中出现了,从而拒绝原假设 H_0 . 原因是,如果原假设 H_0 为真,那么该事件就一定是小概率事件,根据小概率事件在一次实验中认为不可能发生的实际推断原理,我们就应该拒绝 H_0 . 简单地说,假设检验问题就是要在原假设 H_0 和备择假设 H_1 中作出拒绝哪一个和接受哪一个的判断,这类假设检验问题也简称为 H_0 对 H_1 检验问题.

那么根据假设检验的这一原理,在 H_0 对 H_1 的检验问题中要作出某种判断,就必须先抽取一个子样 $(\zeta_1, \zeta_2, \dots, \zeta_n)$, 制定一个法则, 在一定的显著水平 α 下, 一旦子样 $(\zeta_1, \zeta_2, \dots, \zeta_n)$ 的观察值点 $(x_1, x_2, \dots, x_n) \in A$, 我们就拒绝原假设 H_0 , 接受备择假设 H_1 , 称子集 A 为检验的拒绝域(或临界域); 反之, 若子样 $(\zeta_1, \zeta_2, \dots, \zeta_n)$ 的观察值点 $(x_1, x_2, \dots, x_n) \in A^*$, 我们将接受原假设 H_0 , 拒绝备择假设 H_1 , 称子集 A^* 为检验的接受域.

3 假设检验的结果

假设检验的依据是小概率事件在一次实验中认为不可能发生的实际推断原理,但是由于子样的随机性和子集选取的多样性,我们在进行判断时就有可能犯两种错误^[3].

一种错误是,当原假设 H_0 为真时,而子样的观察值点 $(x_1, x_2, \dots, x_n) \in A$, 依照检验法则,我们应当在检验过程中拒绝原假设 H_0 , 这种错误称为第一类错误,即拒真错误. 通常把犯第一类错误的概率记为 α , 表示为 $P(\text{犯第一类错误}) = P(\text{拒绝 } H_0 \mid H_0 \text{ 为真}) = \alpha$, α 可以根据不同问题或具体情况的需要来设置.

另一种错误是,当备择假设 H_1 为真时,而子样的观察值点 $(x_1, x_2, \dots, x_n) \in A^*$, 依照给定的检验法则,我们应当在检验过程中接受 H_0 , 这种错误称为第二类错误,即受伪错误. 通常把犯第二类错误的概率记为 β , 表示为 $P(\text{犯第二类错误}) = P(\text{接收 } H_0 \mid H_1 \text{ 为真}) = \beta$.

对给定的一对对立假设 H_0 和 H_1 , 可以找出许多拒绝域,我们希望所找到的拒绝域 A 可以使犯两类错误的概率都很小. 但是,由于在子样容量 n 一定时,要使 α 和 β 都小是不可能的,只有当子样容量 n 无限增大时才能使犯两类错误的概率都减少^[4], 但这又是不实际的. 因此我们常常以对犯第一类错误的概率 α 进行控制,使得犯第二类错误的概率 β 小为原则,因为相对于错误的接受 H_0 , 人们把拒绝 H_0 看得更重些. 我们把这种只限制犯第一类错误的概率 α , 而不考虑犯第二类错误的概率 β 的统计假设检验问题称为显著性检验问题,称 α 为显著性水平.

4 原假设和备择假设的选取

在两个对立假设中,原假设 H_0 和备择假设 H_1 的选定一般从以下几方面考虑: 首先,要根据问题的具体目的和要求来制定,假如我们希望从子样的观察值中得到对某一观点的支持,那么就把这一观点的否定作为原假设 H_0 , 而把观点本身作为备择假设 H_1 ; 其次,原假设 H_0 是经过严密的调查和取证的,拒绝时应谨慎,有时还要考虑到数学上计算处理时是否方便; 最后^[5], 在一些特殊情况下,要把有可能犯的造成严重后果的错误选定为第一类错误,因为犯第一类错误的概率是我们可以用 α 的大小来控制的,而犯第二类错误的概率 β 是无法控制的.

5 实例应用

下面应用上述理论解决建立在等级基础上的非参数检验问题,以单向等级方差分析法和双向等级方差分析法为例.

5.1 单向等级方差分析

单向等级方差分析是检验 k 个独立的总体样本是否来自同一个总体,具体计算过程如下:

- (1) 抽取样本容量分别为 $n_1, n_2, \cdots, n_k$.
- (2) 选定假设 H_0 : k 个总体无差异(k 个样本来自于同一总体), H_1 : k 个总体有差异(k 个样本来自于不同个总体).
- (3) 确定统计量:
- ①把 $n_1 + n_2 + \cdots + n_k = n$ 个样本由小到大排序,并赋等级分(注意: 相同分数的用平均数代替).

②计算 n_i 样本的等级和 $R_i, H = \frac{12}{n(n+2)} \sum_{i=1}^k \frac{1}{n_i} \left[R_i - \frac{n_i(n+1)}{2} \right]^2$, 如果 $n_1, n_2, \cdots, n_k$ 中有相同的, 则修正为 $H^* = \frac{H}{1 - \frac{\sum_{i=1}^k t_i^3 - t_i}{n^3 - n}} \sim \chi^2(k-1)$ (当 $n \geq 5$ 时), 其中 t_i 是第 i 个群相同的等级中具有相同等级分的观测值的个数.

- (4) 对给定的作出结论,查表 $\chi^2_{\alpha}(k-1)$, 若 $H < \chi^2_{\alpha}(k-1)$ 或 $H^* < \chi^2_{\alpha}(k-1)$, 接受 H_0 ; 否则拒绝 H_0 .
- 例 1 技术专家小组对市场上销售的三种不同品牌(甲、乙、丙) 的电视机进行抽样检测,评分如表 1, 试问这三种不同品牌的电视机质量有无明显差异($\alpha = 0.05$) ?

表 1 甲、乙、丙三种不同品牌的电视机的抽样评分

甲	乙	丙
90	91	97
93	93	99
95	95	99
93.5	92	94
92.5	94	96

- 解: (1) H_0 : 三种品牌的电视机质量无差异; H_1 : 三种品牌的电视机质量有差异.
- (2) 对表 1 中的数据从小到大进行排序,并赋等级分,得出表 2:

表 2 样本的等级分

统一排序	90	91	92	92.5	93	93	93.5	94	94	95	95	96	97	98	99
等级分	1	2	3	4	5.5	5.5	7	8.5	8.5	10.5	10.5	12	13	14	15

整理后得到表 3:

表 3 甲、乙、丙样本的等级和

甲	乙	丙
1	2	13
5.5	5.5	15
10.5	10.5	14
7	3	8.5
4	8.5	12
$R_1 = 28$	$R_2 = 29.5$	$R_3 = 62.5$

- (3) 由表 3 计算如下: $n = n_1 + n_2 + n_3 = 15, H = \frac{12}{n(n+2)} \sum_{i=1}^k \frac{1}{n_i} \left[R_i - \frac{n_i(n+1)}{2} \right]^2$, 代入数据得
- $$H = \frac{12}{15 \times 17} \left[\frac{1}{5} \left(28 - \frac{5 \times 16}{2} \right)^2 + \frac{1}{5} \left(29.5 - \frac{5 \times 16}{2} \right)^2 + \frac{1}{5} \left(62.5 - \frac{5 \times 16}{2} \right)^2 \right] = 7.157$$
- 由于 $t_i = 2$, 所以 H 修正为 $H^* = \frac{7.157}{1 - \frac{6+6+6}{15^3 - 15}} = 7.195$.
- (4) 经查表得: $\chi^2_{0.05}(3-1) = \chi^2_{0.05}(2) = 0.5991, H^* > \chi^2_{0.05}$, 故拒绝 H_0 , 说明三种不同品牌的电视机质量有

显著差异.

5.2 双向等级方差分析

双向等级方差分析是检验 k 个样本是否来自同一个总体,具体计算过程如下:

(1) 把每个样本的 k 种处理方法,从小到大排序并赋等级分,形式如表 4,注意: 若来自同一份总体 R_i 之间差异很小,否则差异很大.

表 4 样本的等级和

	1	2		k
1	X_{11}	X_{12}		X_{1k}
2	X_{21}	X_{22}		X_{2k}
$\vdots$	$\vdots$	$\vdots$		$\vdots$
n	n_{n1}	X_{n2}		X_{nk}
等级和	R_1	R_2		R_k

(2) 选定假设 H_0 : k 种处理方法无差异(k 个样本来自于同一总体) , H_1 : k 种处理方法有差异(k 个样本来自于不同个总体) .

(3) 确定统计量:
$$\chi^2_F = \frac{12}{nk(n+k)} \sum_{j=1}^k R_j^2 - 3n(k+1) \sim \chi^2(k-1).$$

(4) 对给定的作出结论,查表 $\chi^2_{\alpha}(k-1)$,若 $\chi^2_F < \chi^2_{\alpha}(k-1)$,接受 H_0 ; 否则拒绝 H_0 .

例 2 有 12 名药学专家对治疗某种疾病的 5 种药品(A, B, C, D, E) 的主要成分进行测评,评分结果如表 5,试问这 5 种药品的治疗效果有无明显差异($\alpha=0.05$) ?

表 5 5 种药品样本的评分结果

	A	B	C	D	E
1	9	8	7	4	5
2	7	5	8	4	6
3	8	9	7	6	3
4	5	7	9	6	8
5	9	7	5	3	6
6	8	4	9	3	6
7	7	8	5	9	6
8	5	7	6	8	3
9	6	5	7	8	4
10	5	7	6	4	8
11	7	5	8	9	6
12	8	9	6	7	5

解: (1) H_0 : 5 种药品治疗效果无差异; H_1 : 5 种药品治疗效果有差异.

(2) 对表 5 中的数据从小到大进行排序,并赋等级分,得出表 6:

表 6 5 种药品样本的等级和

	A	B	C	D	E
1	5	4	2	1	3
2	4	2	5	1	3
3	4	5	3	2	1
4	1	3	5	2	4
5	5	4	2	1	3
6	4	2	5	1	3
7	3	4	1	5	2
8	2	4	3	5	1
9	3	2	4	5	1
10	2	4	3	1	5
11	3	1	4	5	2
12	4	5	2	3	1
R_i	40	40	39	32	29

- (3) 由表 6 计算如下: $k=5,n=12,\chi_F^2=\frac{12}{nk(n+k)}\sum_{j=1}^kR_j^2-3n(k+1)$ 代入数据得
- $$\chi_F^2=\frac{12}{12\times5\times6}(40^2+40^2+39^2+29^2)-3\times12\times6=3.533$$
- (4) 经查表得 $\chi_\alpha^2(4)=9.488,\chi_F^2<\chi_\alpha^2(4)$,故 5 种药品治疗效果无显著差异.

参考文献:

(1)魏宗舒. 概率论与数理统计教程 (M). 北京: 高等教育出版社, 2008.
(2)缪铨生. 概率与数理统计 (M). 上海: 华东师范大学出版社, 1997.
(3)朱松涛,宋子兴. 概率论与数理统计教程 (M). 济南: 山东大学出版社, 1997.
(4)宋子兴. 关于“假设检验”的原假设与备择假设的讨论 (J). 滨州师专学报, 2002,18(4):33-35.
(5)王嘉澜. 数理统计中关于假设检验的几个要点问题 (J). 高等理科教育, 2005,59(1):79-81.

Hypothetical testing principles and application

WANG Zhi - fu, PAN Xu, JIN Shu, TIAN Feng, WANG Yan

(College of Mathematics and Physics, Bohai University, Jinzhou 121013, China)

Abstract: Hypothetical testing is an important content of mathematical statistics. This article is introduced the principles and the results of hypothesis testing, briefly expounded the setting principles of original hypothesis and alternative hypothesis, applied one - way rank variance analysis and two - way rank variance analysis two kinds of hypothesis testing methods to solve the nonparametric test problems, and simplified the calculation process.

Key words: hypothesis testing; nonparametric test; one - way rank variance analysis; two - way rank variance analysis