

深入理解假设检验原理

康铁祥

(上海财经大学 统计学系,上海 200083)

摘要: 本文通过一道假设检验问题引入了对假设检验原理的思考。文中作者对假设检验中的两类错误及原假设与备择假设之间的关系进行了深入的分析。

关键词: 假设检验 小概率原理 原假设和备择假设 两类错误

中图分类号: O212.1

文献标识码: B

文章编号: 1005-5762(2001)01-0022-02

一、引言

假设检验,尤其是参数假设检验,是现代统计学中必不可少的一部分。但在统计学的研究中,笔者却发现很多统计工作者不能完全理解参数检验的原理。突出表现在对原假设 H_0 和备择假设 H_1 的设定上(尤其是单侧检验)。

例: 某工厂随机选取的 20 只部件的装配时间(分): 9.8, 10.4, 10.6, 9.6, 9.6, 9.8, 10.4, 11.1, 9.6, 10.2, 10.3, 9.6, 9.9, 11.0, 10.5, 9.8, 10.5, 10.1, 10.5, 9.7。设装配时间的总体服从正态分布。问是否可以认为装配时间的均值显著地大于 10(取 $\alpha = 0.05$)?

本例属于假设检验中的单侧检验问题,通常的解法如下:

$H_0: \bar{x} > 10$ $H_1: \bar{x} \leq 10$, 样本量少于 30, 故用 t 检验
又 $S = 0.4696$ $\bar{x} = 10.15$

$$\text{得, } t = \frac{\frac{\bar{x} - 10}{s}}{\frac{1}{\sqrt{20}}} = 1.428 < t_{0.95}(19) = 1.73$$

故接受 H_0 , 即可以认为装配时间的均值显著地大于 10。

那么以上解法究竟正确与否?

二、假设检验理论详述

为了解决这个问题,我们必须首先澄清参数检验的基本原理。参数检验的基本原理是小概率原理,即 如果一个事件发生的概率很小,那么,它在仅有的一次实验中是不可能发生的,并且这一发生的事件应是我们期望的结果。根据这一原理,针对总体中抽取的子样的某些信息,在一定的显著水平情况 α 下,通过设置适当的假设条件 (H_0 与 H_1) 后,我们便可以对总体的参数进行检验。由于样本的取得是随机的,因此, α 不可能为 0, 这样我们就会存在犯错误的可能。我们定义,第 I 类错误: 原假设 H_0 为真, 在检验过程中被拒绝, 即弃真错误; 第 II

除不规则要素, 移动平均后序列记为 $\hat{M}_t^{(D)}$ (关于如何进行 MCD 项移动平均请参阅有关 $X-11$ 法的文献)。由于 $3 \leq \text{MCD} \leq 6$, 所以进行 MCD 项移动平均, 两端缺项不多, 很容易补齐。在第二步已确定的 $\hat{M}_t^{(D)}$ 曲线上的转折点的基础上, 确定曲线 $\hat{M}_t^{(D)}$ 上的对应转折点。选取转折点的方法与第二步类似, 并将纪录转折点序号 $t_i (i = 1, 2, \dots, m)$ 更新为曲线 $\hat{M}_t^{(D)}$ 上的转折点的序号。

第四步, 确定原序列 Y 的转折点。在第三步确定 $\hat{M}_t^{(D)}$ 曲线上的转折点的基础上, 在用类似方法确定原序列 Y 的对应转折点。这样就实现了从光滑序列的转折点逐渐逼近到原序列的转折点, 而原序列上的转折点即

为最终确定的转折点。显见, 时序数据列上转折点的确定, 使用的方法仍为移动平均法。因此, 移动平均法具有测定转折点的功能。

综上所述, 一个看起来似乎很简单的移动平均法, 由于对它的理解层次不同, 切入的视角不同, 讲授的内容、深度和广度不同, 给学生带去的信息量会相去甚远。作为一名高校的教师, 要想讲好所承担课程的每章每节, 都需要下许多功夫, 只有从严治学, 才能不断地提高教学质量。

类错误:备择假设 H_1 正确,在检验过程中却舍弃 H_1 而选择了原假设,称为取伪错误。(见表 1)

表 1 假设检验中的两类错误表

		总 体	
		H_0 为真	H_1 为真
子 样	$\in$ 临界域 (拒绝 H_0)	犯第 I 类错误	正确
	$\notin$ 临界域 (接受 H_0)	正确	犯第 II 类错误

显然 $P(\text{犯第 I 类错误}) = P(\text{拒绝 } H_0 | H_0 \text{ 为真}) = \alpha$, 若令犯第 II 类错误的概率为 β , 则有 $P(\text{犯第 II 类错误}) = P(\text{接受 } H_0 | H_1 \text{ 为真}) = \beta$ 。由数理论的知识,我们可以得到 α 与 β 的一般关系如下:

设子样 $(\xi_1, \xi_2, \dots, \xi_n)$ 取自正态总体 $N(\mu, \sigma^2)$, σ^0 已知,且 n 足够大,对检验假设:

$H_0: \mu = \mu_0, H_1: \mu \geq \mu_0$ 的问题,取临界域 $C = \{(\xi_1, \xi_2, \dots, \xi_n) | \bar{\xi} > c_0\}$, 据此,若犯第 I 类错误的概率为 α 时,可以求出 β (即犯第 II 类错误的概率)。

在 H_0 成立的条件下, $\bar{\xi} \sim N(\mu_0, \frac{\sigma_0^2}{n})$, 则

$$P(\text{犯第 I 类错误}) = \alpha = P(\bar{\xi} \geq c_0) \\ = P\left(\frac{\bar{\xi} - \mu_0}{\sigma_0} \sqrt{n} \geq \frac{c_0 - \mu_0}{\sigma_0} \sqrt{n}\right)$$

$$\text{故 } \frac{c_0 - \mu_0}{\sigma_0} \sqrt{n} = z_{1-\alpha},$$

$$\text{得到 } c_0 = \frac{\sigma_0}{\sqrt{n}} z_{1-\alpha} + \mu_0 \dots \dots (1)$$

其中 $z_{1-\alpha}$ 是 $N(0,1)$ 分布的 $1-\alpha$ 分位点。

在 H_1 成立的条件下, $\bar{\xi} \sim N(\mu, \frac{\sigma_0^2}{n})$, 则

$$P(\text{犯第 II 类错误}) = \beta = P(\bar{\xi} < c_0) \\ = P\left(\frac{\bar{\xi} - \mu}{\sigma_0} \sqrt{n} < \frac{c_0 - \mu}{\sigma_0} \sqrt{n}\right) = \Phi\left(\frac{c_0 - \mu}{\sigma_0} \sqrt{n}\right)$$

将(1)式代入,得到:

$$\Phi\left(\frac{c_0 - \mu}{\sigma_0} \sqrt{n}\right) = \Phi\left(\frac{\frac{\sigma_0}{\sqrt{n}} z_{1-\alpha} + \mu_0 - \mu}{\sigma_0} \sqrt{n}\right) \\ = \Phi\left(z_{1-\alpha} - \frac{\mu - \mu_0}{\sigma_0} \sqrt{n}\right)$$

显然,当 α 增加时, $z_{1-\alpha}$ 减小,从而 β 减少,反之当 α 减小时,将导致 β 的增加。也就是说,我们在实际操作中根本无法找到一个能使 α 与 β 同时减小的临界域,除非增大抽样容量 n ,但是无限增大样本容量 n 不是抽样的本意,因此,本文不予讨论。

为了摆脱假设检验中这一尴尬的局面,奈曼与皮尔逊(Neyman & Pearson) 提出一个原则,即在控制犯第 I 类错误的概率 α 的条件下,尽量使犯第 II 类错误的概率

β 小。这其中蕴含了这样一种思想:我们只对犯第 I 类错误的概率 α 加以限制,而不考虑犯第 II 类错误的概率。根据这一原则,我们在假设检验中更倾向于拒绝 H_0 而不是接受 H_0 。换句话说,如果我们拒绝了 H_0 ,我们就有 $1-\alpha$ 把握相信 H_0 为伪;但若接受 H_0 ,我们只能说没有足够的证据证明 H_0 为伪,仅此而已。又由小概率原理知,如果某一事件在仅有的一次实验中发生了,那我们初步断定事件不应属于小概率事件,并且也有充分的理由相信检验结果应如实验所反映的那样,因此,在实际应用中,应将实验中发生的反现象设为原假设 H_0 。我们是这样考虑的:对实验结果的反现象进行让步,如果在此条件下,得出了拒绝原假设 (H_0) 的结论,那我们就完全有理由接受 H_1 ;相反,我们只能说没有足够的证据证明 H_0 是错误的。

三、结论

由上所述,本文开头例题正确的分析及解法如下:

由于在一次抽取中我们得到了 $\bar{x} = 10.15 > 10$, 故我们认为 $10.15 > 10$ 这一事件不应属于小概率事件,故假设 $H_0: \bar{x} \leq 10, H_1: \bar{x} > 10$, 且有,

$$S = 0.4696, \bar{x} = 10.15, \text{样本量少于 } 30, \text{故用 } t \text{ 检验,} \\ \text{得, } t = \frac{\bar{x} - 10}{\frac{s}{\sqrt{20}}} = 1.428 < t_{0.95}(19), \text{因此我们没有}$$

充分的理由拒绝 H_0 。

其实,现在这道题还没有完全完成。这是因为根据奈曼与皮尔逊原则,在假设检验过程中我们未对犯 II 类错误的概率 β 进行控制,这样,若实验过程中接受了 H_0 ,我们就无法确定是否犯了第 II 类错误。所以,为了减少犯第 II 类错误的概率,笔者认为应尽量进行二次抽样,即,由于我们控制了犯第 I 类错误的概率 α , 故应注意犯第 II 类错误的概率 β , 进行二次抽样的目的是将犯第 II 类错误的概率减少为 β 。如果二次抽样的结果依然是接受 H_0 ,我们就有一定的理由接受 H_0 或加大抽样容量进行重新抽样。

通过以上的分析,在参数检验中我们可以得到这样的结论:

1. 参数检验依靠的第一思想是小概率原理,但在一次实验中发生的事件不再是小概率事件;
2. 备择假设 H_0 是我们进行了让步的事件,它应该是实验结果的反事件,也就是我们认为的小概率事件;
3. 我们更希望的检验结果是拒绝原假设 (H_0),而不是接受原假设。