8.7 检验的p值

先看一个假设检验问题.

例5.1.2 为来自总体的简单随机样本,那么检验问题

 对 

的水平为的拒绝域为,水平为的拒绝域为,其中检验统计量.如果由样本值得到样本均值为,那么检验统计量的观察值为,从而在检验水平下,可得出拒绝原假设的结论,而在检验水平下,得不出拒绝原假设的结论.这种现象会给我们在实际的检验中带来不便.

我们继续讨论以上问题,如果,那么在水平下我们得出拒绝原假设的结论;如果,在水平下我们也是得出拒绝原假设的结论.这两种情况下我们得出相同的结论.但在后一种情况下我们得出拒绝原假设的结论的“底气”似乎更足.如果只是在给定的检验水平下在“接受原假设”和“拒绝原假设”之中作出二选一的判断.这样的判断的依据有多大?或者说这个判断的“底气”有多足?则往往不能清晰地显示出来.

为了在实际应用中克服以上种种缺陷,人们引入了“检验的值”.

定义5.1.5 在一个具体的检验问题中,依据观察到的结果,能得出拒绝原假设之结论的最小的检验水平称为检验的值.

续例5.1.2

如果,即,由于,那么当检验水平不小于时,检验结论是拒绝原假设;当检验水平小于时,检验结论是不拒绝原假设,因此此时的值为.又如,即,由于,此时的值为.

由此可以看出,在由样本值计算出值后,如果检验水平值,则结论为拒绝原假设;如果检验水平值,则结论为不拒绝原假设.

在实际问题中, 检验水平给多大合适并无一矩之规.因此在实用中我们常常不预先指定检验水平，而是直接计算出值.值越小，拒绝原假设的证据越足. 反之若值越大,则认可原假设的证据也越充足.

检验的值的意义可作如下解释:考虑检验问题.如果该检验问题的拒绝域的形式为,这就意味着检验统计量的值越大，样本与原假设的偏离程度越大,或者说样本与原假设的矛盾越大.如果由具体的样本值得到的观察值为,那么概率表示了出现如此“偏离” (相对于原假设的偏离)程度的可能性有多大.若这个概率不小，说明出现这种程度的“偏离”属正常，无须大惊小怪,因而没有理由拒绝原假设.反之, 若这个概率很小,说明出现这种程度的“偏离”几乎是不可能的,太不正常了.由于这个概率是在原假设为真的条件下得出的,因而这种“太不正常”的现实只是意味着原假设对应的随机机制只乎不可能产生“如此的样本数据”,所以有理由拒绝原假设，而且概率越小拒绝原假设的理由越充分. 检验的值可以作为度量样本与原假设之的矛盾程度的指标，值越小,矛盾越大.

那么值是怎么算出来呢？由例5.1.2，我们已经能够看出来。一般地，若检验问题的拒绝域的形式为（这里说“形式”的意思是我们无需确定），其中为检验统计量.在有了样本值后计算出检验统计量的观察值为，那么的值为.若检验问题的拒绝域的形式为或.检验统计量的观察值为，那么的值为,由此可以看出，为计算检验的值，需要计算检验统计量的的观察值，以及确定检验统计量的零分布.在具体问题中,只要有了具体的样本值计算检验统计量的的观察值并不困难,而确定检验统计量的零分布往往不易,常需求助于近似的零分布或用随机模拟方法.

例 续例5.1.2

（i）检验统计量的观察值为,而.即检验的为0.0013,它小于检验水平,故拒绝原假设.

(ii) (1) 检验统计量的观察值为,而.即检验的为超过0.025,它大于检验水平,故接受原假设.

(2). 检验统计量的观察值为,而.即检验的值小于检验水平,故拒绝原假设.

p值在实际中很有用,如今的统计软件中对检验问题一般都会给出检验的p值.

我们在以后的检验可以从两方面进行,其一是建立拒绝域,考察样本观测值是否落入拒绝域而作出判断;其二是根据样本观测值计算检验的p值,再根据检验的p值的大小作出判断.