

## 假设检验

在总体的分布函数完全未知或只知其形式,但不知其参数的情况



为了推断总体的某些未知特性



提出某些关于总体的假设 (如:提出总体服从泊松分布的假设,又如:对于正态总体提出数学期望等于 $\mu_0$ 的假设等)



根据样本对所提出的假设作出是接受,还是拒绝的决策

既假设总体服从某种性质,通过样本检验是否接受该假设

例如:假设  $H_0: \mu = \mu_0 = 0.5$  其中  $\mu$  为总体均值,假设其等于 0.5, 总体标准差  $\sigma$  已知。  
 $H_1: \mu \neq \mu_0$

① 目标: 要检验  $\mu$  是否等于  $\mu_0$ , 想到借用样本均值  $\bar{x}$  来判断

∵  $\bar{x}$  是  $\mu$  的无偏估计

分析: 若  $H_0$  为真, 则  $\bar{x}$  与  $\mu_0$  的偏差应该不会太大, 若

∵  $\bar{x}$  的观察值  $\bar{x}$  的大小一定程度反映  $\mu$  的大小

$|\bar{x} - \mu_0|$  太大, 则怀疑  $H_0$  的正确性从而拒绝  $H_0$ 。



② 目标: 判断  $|\bar{x} - \mu_0|$  是否太大

分析: 若  $H_0$  为真, 则  $\frac{\bar{x} - \mu_0}{\sigma/\sqrt{n}} \sim N(0, 1)$ , 可以用  $\frac{|\bar{x} - \mu_0|}{\sigma/\sqrt{n}}$  的大小衡量  $|\bar{x} - \mu_0|$  的大小。

希望有一个阈值  $k$  当  $\frac{|\bar{x} - \mu_0|}{\sigma/\sqrt{n}} > k$  时表明  $|\bar{x} - \mu_0|$  太大, 既  $\bar{x}$  与  $\mu_0$  的偏差太大从而拒绝  $H_0$ 。



③ 目标: 寻找合适的阈值  $k$

分析: 由于此处使用的是样本均值, 即使  $H_0$  为真, 也可能出现  $|\bar{x} - \mu_0|$  过大, 从而判断  $H_0$  为假

此时就出现了错误既  $H_0$  为真但检验后却拒绝  $H_0$ , 而且这种错误还想不到方法消除

因为该错误无法消除, 所以希望控制在一定范围内

目标①: 希望将错误的概率控制在一个较小的数  $\alpha$  内即  $P\{\text{当 } H_0 \text{ 为真拒绝 } H_0\} \leq \alpha$

即允许犯错的概率最大为  $\alpha$

即接受在有  $\alpha$  概率出错的情况下仍然愿意相信检验结果

也就是说在满足  $P\{\text{当 } H_0 \text{ 为真拒绝 } H_0\} \leq \alpha$  时能接受检验结果

分析: ① 你能接受的错误概率  $\alpha$  越大, 则  $k$  的取值会越小, 此时  $|\bar{x} - \mu_0|$  越容易过大, 越容易拒绝

②  $\alpha$  越大说明当  $H_0$  为真但拒绝  $H_0$  的概率越大, 说明检验的结果拒绝的概率越大, 也说明拒绝域越大

由以上分析,寻找阈值 $k$ ,可以通过满足  $P\{\text{当 } H_0 \text{ 为真拒绝 } H_0\} = P_{\mu=\mu_0}\left\{\left|\frac{\bar{x}-\mu_0}{\sigma/\sqrt{n}}\right| > k\right\} \leq \alpha$  约束,通过该式就可以确定 $k$ ,当 $k$ 越小时  $P_{\mu=\mu_0}\left\{\left|\frac{\bar{x}-\mu_0}{\sigma/\sqrt{n}}\right| > k\right\}$  越大.  
 $\therefore$  由  $P_{\mu=\mu_0}\left\{\left|\frac{\bar{x}-\mu_0}{\sigma/\sqrt{n}}\right| \geq k\right\} = \alpha$  确定

以上为双边假设检验,单边检验理解类似,参看例子教大数据大作业第四题推导

$\therefore$  处理参数的假设检验问题的步骤如下

- ① 根据实际问题要求,提出原假设 $H_0$ 及备择假设 $H_1$
- ② 给定显著性水平 $\alpha$ 以及样本容量 $n$
- ③ 确定检验统计量以及拒绝域的形式
- ④ 按  $P\{\text{当 } H_0 \text{ 为真拒绝 } H_0\} \leq \alpha$  求出拒绝域
- ⑤ 取样,根据样本观察值作出决策是接受 $H_0$ 还是拒绝 $H_0$ .