假设检验

Guangyao Zhao

2022-09-01

Contents

个人认为假设检验是概率论与数理统计最重要的一章，同时也是在非数学专业的工科中应用最广泛的一章。在我自己的研究中也经常会应用到假设检验理论，尤其是在判断相关性是否成立的时候。

原理

假设检验的基本原理：小概率事件在一次实验中几乎不发生。这是一个非常有意义的假设，因为概率论本身基于的就是大家公认的常识，在生活中通常我们也会认为小概率事件不发生。

构造某事件 A，在某假设 H_0 为真的条件下，H_1 发生概率的就会很小，然后根据样本观察值得知这个小概率事件，即 H_1 是否发生了。如果发生了，就拒绝 H_0，如果没发生，就接受 H_0。

比如对于事件 A 为『今天是否会下雨』，假设 H_0 为『今天会下雨』为真的条件下，对立面的小概率事件 H_1 就为『今天不下雨』，如果小概率事件 H_1 真的发生了，那就说明原假设 H_0 错误，即拒绝 H_0。

假设检验的流程如下：

1. 假设 H_0 为真
2. 构造统计量
3. H_0 的对立面 H_1 发生概率很小
4. 如果 H_1 发生，则拒绝 H_0，反之则接受

显著水平

一批机器正常时，均值为 $0.5\,kg$，对一批样本是否正常进行假设检验：

- 机器正常 (原假设)：$H_0 : \mu = \mu_0 = 0.5$
- 机器不正常 (备择假设)：$H_1 : \mu \neq \mu_0$

在假设检验中，不可避免的可能发生以下两类错误：

- H_0 为真，拒绝 H_0
- H_0 为假，接受 H_0

既然错误无法排除，那么只能控制发生错误的概率，一般情况下只考虑第一类错误发生的概率情况 (称之为显著性检验)，即：

$$P\{H_0\text{为真}, \text{拒绝}H_0\} \leq \alpha$$

$P\{H_0\text{为真}, \text{拒绝}H_0\}$ 的意思是在假设 H_0 为真的前提下，拒绝 H_0 的概率，即小概率事件 H_1 的概率 $\leq \alpha$ 的时候，拒绝 H_0。上述 α 称之为显著水平。

具体一点的解释，在原假设 H_0 为真的前提下，小概率事件即备择假设 H_1 发生了，依据常识，小概率事件发生的概率特别特别小，但是一次试验竟然发生了，那么就可以认为原假设是错误的，即原假设为假，备择假设为真。显著水平 α 常常选取 0.05，即小概率事件小到 5% 的时候竟然发生了，那就索性说原假设为假了。

假设检验的基本概念

原假设和备择假设的选择标准

原假设和备择假设原则上可以随便假设，但是应该尽量遵守以下规则：

- 应该把大众普遍认为成立的命题作为原假设：原假设不能轻易拒绝，比如长跑对健康有益是常识，应该作为原假设。
- 应该把分析人员想证明不正确的命题作为原假设：据说有种新方法能改进生产效率。我们的目的是证明它是真的，所以应该把『新方法能改进生产效率』作为备择假设。

检验统计量

根据假设和条件确定的一个统计量，并在 H_0 成立的条件下可以确认其分布

拒绝域和临界点

- 某个可以拒绝 H_0 的区域称之为拒绝域
- 拒绝与的边界点称之为临界点

例子 1

因为 $\bar{X}$ 是 $\mu=0.5$ 的无偏估计，若 H_0 为真，此时 $|\bar{X}-0.5|$ 应该比较小，即 $\bar{X}$ 应该落在 μ 周围。因此在 H_0 为真的条件下，拒绝 H_0 应该满足 $|\bar{X}-0.5|$ 比较大。H_0 为真的条件

$$Z=\frac{\bar{X}-0.5}{\sigma/\sqrt{n}} \sim N(0,1)$$

此处的 Z 称之为检验统计量（枢轴量），其中 $\bar{X},\sigma,n$ 都是已知量。可以认为当 $\frac{\bar{X}-0.5}{\sigma/\sqrt{n}}$ 较大（小概率事件发生）时，可以作出拒绝 H_0 的结论，即当 $\frac{\bar{X}-0.5}{\sigma/\sqrt{n}}>k$ 时，拒绝 H_0，而常数概率 k 由

$$P\{\text{拒绝}H_0|H_0\text{为真}\}=P\left\{\frac{\bar{X}-0.5}{\sigma/\sqrt{n}}>k\right\}=\alpha$$

可得：$k=z_{\alpha/2}$

故当统计量的观察值满足

$$|z|=\frac{\bar{x}-0.05}{\sigma/\sqrt{n}}>z_{\alpha/2}$$

时，拒绝 H_0。

当 $|z|>z_{\alpha/2}$ 时，对照正态分布的概率图形可知，$|\bar{X}-0.05|$ 中心点太远了，即 $|\bar{X}-0.05|$『应该比较小』这件事不足够小了。

在刚开始学习假设检验的时候我一直没注意对于假设检验，标准正态分布也好、t 分布也好等等分布的图形的含义。其实可以将其横轴理解为估计量的分布（实际上在学习分布概率的时候也是这么学习的），真实值居中（如果是对称分布），大多数的估计值应该接近真实值，只有少部分才会偏离，当估计值偏离到一定程度时，就可以认为该估计值是真实值的概率就很小了，即**假设的**小概率事件发生了，进而也就可以拒绝原假设了。

例子 2

假设事件 A 为『我有一个特别的打电话技巧，接我电话的都是女生』。此时可做出假设：

- H_0 (原假设)：无技巧
- H_1 (备择假设)：有技巧

做实验：打了 20 个电话，其中有 18 个是女生回复的。那么该事件发生的概率为：

$$P = \left(\frac{1}{2}\right)^{20} C_{20}^{18} + \left(\frac{1}{2}\right)^{20} C_{20}^{19} + \left(\frac{1}{2}\right)^{20} C_{20}^{20} = 0.0002$$

这个概率足够小了吧，但是竟然发生了。那么，概率多小算小呢？此处的 P-value 又是什么呢？

- 显著水平：一般情况下的参考标准是 $1 - \alpha = 0.95$，即在原假设正确的前提下，备择假设发生的概率 ≤ 0.005 时则可拒绝原假设
- P-value：0.0002，即在在假设原假设正确时，出现现状或更差的情况的概率。此处的现状指的是『打了 20 个电话，其中有 18 个是女生回复的』，更差的情况指的是『打了 20 个电话，其中有 19 或者 20 个是女生回复的』