Lecture 13

假设检验

Probability and Statistics
Beihang University

13.1 参数假设检验

前面我们关心的是参数估计的问题, 但实际中关于参数的问题不仅限于此. 例如某车间生产滚珠, 由于机器性能不同, 致使实际滚珠直径 X 有所差异, 经验表明 $X \sim \mathcal{N}\left(\mu, \sigma^2\right)$, 其中 μ 由设计决定, σ^2 由机器性能决定. 在正常情况下, $\sigma^2 \leqslant \sigma_0^2$. 但由于长时间使用, 机器性能会发生变化. 因此所关心的一个问题是, 机器是否异常, 即是否有 $\sigma^2 > \sigma_0^2$:

$$H_0 : \sigma^2 > \sigma_0^2 \longleftrightarrow H_1 : \sigma^2 \leqslant \sigma_0^2$$

在这个问题中所需要回答的问题, $\sigma^2 > \sigma_0^2$, 称为原假设. $\sigma^2 \leqslant \sigma_0^2$ 则称为其备选假设. 一般地

定义 13.1.1 假设 X 服从某参数为 θ 的分布, 其中 $\theta \in \Theta$. 在一个统计检验过程中, 参数集合 Θ 被划分为两个不相交的部分 $\Theta = \Theta_0 \cup \Theta_1$. 用 H_0 表示 $\theta \in \Theta_0$ 这一假设, 并称之为原假设; H_1 表示 $\theta \in \Theta_1$ 这一假设, 称为备选假设. 如果决定了 $\theta \in \Theta_0$, 那么

我们称拒绝原假设 H_0.

例题 13.1.1 假设 $X_1,\ldots,X_n$ 为来自于 $N\left(\mu,\sigma^2\right)$ 的正态分布. 其中 σ^2 已知. 想要检验

$$H_0:\mu=\mu_0$$

$$H_1:\mu\neq\mu_0$$

- 若 $\overline{X}_n$ 与 μ_0 相距较远, 那么我们可以合理地拒绝 H_0. 因此我们选取 $c>0$, 令 $T=\left|\overline{X}_n-\mu_0\right|$,

$$S_0=\left\{(X_1,\ldots,X_n):T(X_1,\ldots,X_n)\leqslant c\right\},S_1=S_0^c$$

那么当样本 $(X_1,\ldots,X_n)\in S_1$ 时, 即 $T>c$, 拒绝 H_0.

在这一例子中我们看到, 所关心的参数集合为

$$\Theta_0=\{\mu_0\},\Theta_1=\mathbb{R}\setminus\{\mu_0\}$$

由 $(X_1,\ldots,X_n)$ 做成的样本空间划分为 S_0 和 S_1 两个不相交的部分, 我们将依据如下关系做出判断,

$$\mu\in\Theta_0\Longleftrightarrow(X_1,\ldots,X_n)\in S_0$$

$$\mu \in \Theta_1 \Longleftrightarrow (X_1, \ldots, X_n) \in S_1$$

可见,

$$\text{如果}\mu \in \Theta_0, \text{判断出错} \Longrightarrow \text{I 类错误, 概率} P\left(S_1 | \mu\right)$$

$$\text{如果}\mu \in \Theta_1, \text{判断出错} \Longrightarrow \text{II 类错误, 概率} 1 - P\left(S_1 | \mu\right)$$

自然地在做检验时, 希望将两类错误的概率都降到最低, 然后这往往很难同时做到, 因此我们有

定义 13.1.2 一个统计检验的显著水平定义为

$$\alpha_0 = \sup_{\theta \in \Theta_0} P\left(S_1 | \theta\right)$$

显著水平给出了 I 类错误概率的一个上界, 然后 II 类错误却被放在一个相对次要的位置上. 这实际上引入了 I 类错误和 II 类错误的非对称性, 这种合理性来源与对实际问题需要的考虑, 例如, 一个检验中有两类错误：将一个变异的细胞判断为一个

正常细胞, 或者将正常细胞判断为变异细胞, 那么我们自然地认为前一个错误会带来较为严重的后果. 因此我们更希望将前一个错误发生的概率降到最低. 基于此, 常常将较为关注的问题假设作为原假设 H_0.

例题 13.1.2 假设某工厂生产强度指标为 X 的瓦片, 根据经验, $X \sim \mathcal{N}\left(\mu, \sigma_0^2\right)$, $\sigma_0^2 = 1.21$. 现抽出 6 块瓦片测得其强度指标分别为：32.56, 29.66, 31.64, 30.00, 31.87, 30.23. 问在显著水平 $\alpha_0 = 0.05$ 的条件下, 这批瓦片的强度指标是否超过 30.

■ 根据题意, 厂家希望瓦片强度指标达到 $\mu_0 = 30$, 而强度较低的则认为不合格, 不合格产品是厂家较为关注的, 因此, 设待检验问题为

$$H_0 : \mu \leqslant \mu_0 \longleftrightarrow H_1 : \mu > \mu_0$$

由于 σ_0^2 已知, 选取统计量

$$T = \frac{\overline{X}_n - \mu_0}{\sigma_0/\sqrt{n}} \sim \mathcal{N}\left(\mu - \mu_0, 1\right)$$

注意 $E\left(\overline{X}_n\right)=\mu$, 如果 $\overline{X}_n-\mu_0$ 较大, 或者 T 较大, 那么我们应拒绝 H_0. 我们选取 $c>0$ 使得

$$\sup_{\mu\leqslant\mu_0} P\left(T>c|\mu\right)=P\left(T>c|\mu_0\right)=\alpha_0$$

即有

$$P\left(T\leqslant c|\mu_0\right)=1-\alpha_0, c=\Phi^{-1}\left(1-\alpha_0\right)$$

当 $T>c$ 时拒绝 H_0. 在本例中 $\alpha_0=0.05$, $c=1.65$, $n=6$, $T=2.212>1.65$, 因此拒绝原假设 H_0, 即认为该产品是合格的. 此时, 犯一类错误的概率被控制在 0.05 以下, 因此结论有很大的可靠性.

例题 13.1.3 设某厂生产铜丝, 铜丝的质量由其拉断力方差决定, 拉断力方差越大, 铜丝质量越差. 经验表明拉断力 $X\sim\mathcal{N}\left(\mu,\sigma^2\right)$. 厂家认为 $\sigma^2<\sigma_0^2=8$ 为合格. 现抽取一些铜丝并测量得到拉断力如下：578, 512, 570, 568, 572, 570, 570, 572, 596, 564. 问在显著水平 $\alpha_0=0.05$ 的条件下, 铜丝质量是否合格.

■ 厂家比较关注的是不合格的产品, 因此设待检验问题为

$$H_0 : \sigma^2 \geqslant \sigma_0^2 \longleftrightarrow H_1 : \sigma^2 < \sigma_0^2$$

这等价于

$$H_0 : \frac{1}{\sigma^2} \leqslant \frac{1}{\sigma_0^2} \longleftrightarrow H_1 : \frac{1}{\sigma^2} > \frac{1}{\sigma_0^2}$$

考虑统计量

$$T = \frac{1}{\sigma_0^2} \sum_{i=1}^{n} \left(X_i - \overline{X}_n\right)^2 .$$

令 $\chi_\alpha^2(n-1)$ 为 $\chi^2(n-1)$ 对应于 α 的分位数, 那么当 $T < \chi_\alpha^2(n-1)$ 时否定原假设. 在本例中, 取 $\alpha = 0.05$, $n = 10$, $T = 63.65$, 查表得 $\chi_\alpha^2(n-1) = 3.325$, 因此不否定原假设, 即产品需要进一步完善. 由于原假设 H_0 并未被否定, 犯二类错误的概率没有得到控制, 因此结论并不是很强的.