第五章 参数假设检验

§5.1 假设检验的若干基本概念

"反证法"：

给定假设⇒在假设条件下推出矛盾的结果⇒否定假设

在统计中

给定假设⇒在假设下若发生了不太合理的现象⇒否定假设

假设检验：

给定假设⇒在假设下若发生了小概率事件⇒否定假设(否则，则保留假设)

推断理由：实际推断原理

小概率事件在一次实验中是不会发生的

一、什么是假设检验?

Example

某餐厅每天营业额服从正态分布，以往老菜单其均值为7000元，标准差为640元．一个新菜单挂出后,九天中平均每天的营业额为7300元,经理想知道这个差别是否出于菜单而引起的,即,新菜单是否比老菜单好? (这里假定标准差不变).

这个例子中涉及两个对象:

- 老菜单下的每天营业额: $N(\mu_0, \sigma_0^2)$, $\mu_0 = 7000$, $\sigma_0 = 640$;
- 新菜单下的每天营业额: $N(\mu, \sigma_0^2)$, $\sigma_0 = 640$, μ未知.
- 得到的数据是: 新菜单下的一个样本$X_1, \cdots, X_n$的样本均值$\overline{x} = 7300$, 它是μ的一个估计值;
- 要判断的是: 是否有$\mu > \mu_0 = 7000$.

假设检验的做法:

(1)建立假设.

先建立一个命题："新老菜单的平均营业额之间没有差异". 这个命题称为原假设(或称为零假设), 记为H_0. 我们的任务就是确认H_0是真还是假.

当确认H_0为假时就拒绝(抛弃)H_0,这时我们就选择备择假设(或称对立假设): "新菜单下的平均营业额比老菜单的高", 记为H_1.

上述假设检验问题可以表示

$$H_0: \mu = \mu_0 \longleftrightarrow H_1: \mu > \mu_0.$$

当我们拒绝原假设H_0时就可以认为H_1为真.

(2)构造检验统计量

假设检验的做法是:先假定H_0为真,然后用样本判断其真伪.

由于样本所含信息分散, 因此一般需要构造合适的统计量来做判断, 称之为检验统计量.

这里取检验统计量为$\overline{X}$.

在H_0为真时, 总体的分布是$N(\mu_0, \sigma_0^2)$, 从而H_0为真时,

$$\overline{X} \sim N(\mu_0, \sigma_0^2/n) = N(7000, \frac{640^2}{9}).$$

(3) 确定拒绝域与接受域

在H_0为真时，$\overline{X}$的观察值$\overline{x}$应接近7000. 在H_1为真时，$\overline{X}$的观察值$\overline{x}$应该比7000大.

故将样本空间Ω分成两部分:

$$D = \{(x_1, x_2, \cdots, x_n) : \overline{x} > c\}, - - - - -\text{拒绝域}$$
$$\overline{D} = \{(x_1, x_2, \cdots, x_n) : \overline{x} \le c\}, - - - - -\text{接受域}.$$

$D \cup \overline{D} = \Omega.$

其中c称之为检验的临界点.

(4) 据给定的显著性水平来确定临界点

对原假设H_0是否为真作判断时可能会犯错误，这就是要冒风险,这一风险，我们用一个概率表示,这个概率是

$$\mathsf{P}(H_0\text{被拒绝}|H_0\text{为真}) = \mathsf{P}(\overline{X} > c|\mu = \mu_0) = \mathsf{P}_{\mu_0}(\overline{X} > c).$$

我们要求这个概率不超过α (显著性水平). 这里取$\alpha = 0.05$, 并要求这个概率等于α, 那么有

$$\mathsf{P}_{\mu_0}(\overline{X} > c) = 1 - \Phi\left(\frac{c - 7000}{640/3}\right) = 0.05,$$

即得$c = 7000 + 1.645 \times \frac{640}{3} = 7350.9$. 从而拒绝域为

$$D = \{\overline{x} > 7350.9\}.$$

(5) 做判断

在H_0为真的前提下，$\{\overline{X} > c\} = \{(X_1, X_2, \cdots, X_n) \in D\}$ 这一事件发生的概率为0.05，是一个小概率事件. 通常在一次试验中小概率事件是难以发生的,倘若这一事件在一次试验中发生了，人们就有理由怀疑它不是小概率事件. 这一矛盾导致人们不相信原假设H_0为真,从而否定原假设.

所以检验准则为

$$\text{当}\overline{x} > 7350.9\text{时，拒绝}H_0,$$

$$\text{当}\overline{x} \le 7350.9\text{时，保留}H_0.$$

现在$\overline{x} = 7300 < 7350.9$，故应保留$H_0$，即判断:新菜单的挂出对平均每天的营业额没有显著性影响.

二、假设

假设: 指的是关于总体分布的命题.

在一个假设检验问题中,常涉及两个假设:

所要检验的假设称为原假设(或称为零假设),记为H_0.

与H_0不相容的假设称为备择假设(或称为对立假设),记为H_1.

给定了H_0和H_1相当于给定了一个检验问题,有时也记为检验问题(H_0, H_1).

参数假设: 在参数分布族$\mathcal{F}=\{p_\theta(x):\theta\in\Theta\}$场合所作的关于其未知参数的假设. 此时,原假设和备择假设可分别记为

$$H_0:\theta\in\Theta_0,\qquad H_1:\theta\in\Theta_1,$$

其中, Θ_0与Θ_1是Θ的两个互不相交的非空子集.

除了参数假设以外的其它假设称为非参数假设.

对于参数假设, 当Θ_0(或Θ_1)中只含一个元素时, 则称该假设为简单假设, 否则称为复杂假设.

注:

1. 原假设H_0和备择假设H_1不相容,但不一定互补;
2. 原假设H_0是不易被否定的, 原假设H_0和备择假设H_1对于一定立场而言不可以互换;
3. (不成文规定)参数假设检验问题,原假设H_0中大多含有"="。

三、检验

检验：给出一个规则，凭此规则，在有了样本观测值之后，就可以作出接受还是拒绝原假设H_0的判断，这样的规则称为检验.

检验的实质：给出了样本空间Ω的一个分划，即将样本空间Ω分成了两个互斥的集合，

$$\Omega = D \cup \overline{D}.$$

当$\tilde{x} \in D$时，就拒绝H_0，称D为拒绝域(或否定域)；

当$\tilde{x} \in \overline{D}$时，就接受$H_0$，称$\overline{D}$为接受域. 这就是检验的规则.

为了数学上的处理方便，有时也引入**检验函数**$\varphi(\tilde{x})$:

$$\varphi(\tilde{x}) = \begin{cases} 1, & 当\tilde{x} \in D; \\ 0, & 当\tilde{x} \in \overline{D}. \end{cases}$$

检验函数$\varphi(\tilde{x})$ 与检验规则相对应.

为了确定拒绝域, 往往需要构造一个统计量. 称能从样本空间中划分出拒绝域的统计量为检验统计量.

给定一个检验$\Leftrightarrow$ 给定一个拒绝域$\Leftarrow$ 通过检验统计量$T(\tilde{X})$.

四、两类错误与势函数

第一类错误: 原假设$H_0 : \theta \in \Theta_0$为真, 但由于样本的随机性,样本观测值落入拒绝域$D$, 此时所下的判断是拒绝$H_0$, 这类错误称为第一类错误(或称弃真错误), 其发生的概率称为犯第一类错误的概率, 弃真概率, 通常记为$\alpha(\theta)$,

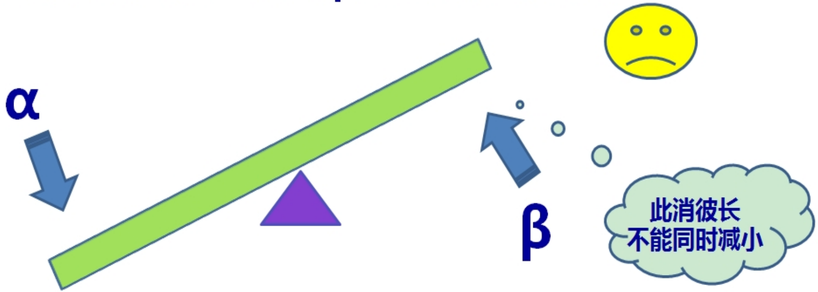
$$\alpha(\theta) = \mathsf{P}(\text{拒绝}H_0|H_0\text{为真}) = \mathsf{P}(\widetilde{X} \in D|H_0\text{为真}) = \mathsf{P}_\theta(\widetilde{X} \in D), \quad \theta \in \Theta_0.$$

第二类错误: 原假设$H_0 : \theta \in \Theta_0$为假, 但由于样本的随机性,样本观测值落入接受域, 这时所下的判断是接受H_0, 这类错误称为第二类错误(或称取伪错误/存伪错误), 其发生的概率称为犯第二类错误的概率, 取伪概率, 通常记为$\beta(\theta)$,

$$\beta(\theta) = \mathsf{P}(\text{接受}H_0|H_1\text{为真}) = \mathsf{P}(\widetilde{X} \in \overline{D}|H_1\text{为真}) = \mathsf{P}_\theta(\widetilde{X} \in \overline{D}), \quad \theta \in \Theta_1.$$

n固定时,α小, β就大; β小, α就大.

第一类错误α 与第二类错误β 的关系（当n固定时）



采用Neyman-Pearson(NP)原则进行“妥协”

设假设检验的拒绝域为D, 则

$$\mathsf{P}_\theta(\widetilde{X} \in D) = \begin{cases} \text{犯第一类错误的概率}, & \text{当}\theta \in \Theta_0, \\ 1 - \text{犯第二类错误的概率}, & \text{当}\theta \in \Theta_1. \end{cases}$$

Definition

定义5.1.2 设检验的拒绝域为D, 称

$$g(\theta) = \mathsf{P}_\theta(\widetilde{X} \in D), \quad \theta \in \Theta$$

为此检验的power function.

五、Neyman-Pearson原则与显著性水平为α的检验.

控制犯第一类错误的概率,即选定一个数α, $0<\alpha<1$, 要求检验犯第一类错误的概率不超过α:

$$\sup_{\theta\in\Theta_0}\alpha(\theta)\leq\alpha.$$

犯第一类错误的概率不超过α的检验,称为显著性水平为α的检验.

通常取检验临界点使得

$$\sup_{\theta\in\Theta_0}\alpha(\theta)=\alpha \text{ 或小于但尽量接近}\alpha.$$

六、处理假设检验问题的一般步骤

1. 根据问题提出原假设H_0和备择假设H_1;
2. 确定检验统计量$T(X_1, X_2, \cdots, X_n)$, 并根据原假设和备择假设确定拒绝域D的形式.(拒绝域D的形式主要依赖于备择假设的形式)

(1) 单侧拒绝域

$$D = \{\tilde{x} : T(\tilde{x}) < C\};\quad \text{或}; D = \{\tilde{x} : T(\tilde{x}) > C\};$$

(2) 双侧拒绝域

$$D = \{\tilde{x} : C_1 < T(\tilde{x}) < C_2\};\quad \text{或}; D = \{\tilde{x} : |T(\tilde{x})| < C\};$$

$$\text{或};\ D = \{\tilde{x} : T(\tilde{x}) < C_1 \ \text{或}\ T(\tilde{x}) > C_2\};$$

$$\text{或};\quad D = \{\tilde{x} : |T(\tilde{x})| > C\};$$

3. 选取适当的显著性α，并求出临界值，使得

$$\sup \mathsf{P}(\widetilde{X} \in D | H_0为真) \leq \alpha并尽可能地接近\alpha.$$

在总体为连续型随机变量时，往往要使得

$$\sup \mathsf{P}(\widetilde{X} \in D | H_0为真) = \alpha.$$

4. 由样本$\widetilde{X}$的观察值$\widetilde{x}$算出检验统计量$T(\widetilde{X})$的值$t = T(\widetilde{x})$，并与临界点进行比较. 若观察值$\widetilde{x}$落入拒绝域D，则拒绝原假设H_0，否则接受原假设.