第八章

假设检验

§8.1 假设检验的基本概念

若对参数一无所知 → 用参数估计的方法处理

若对参数有所了解 → 但有怀疑猜测需要证实之时 → 用假设检验的方法来处理

△ 何为假设检验?

假设检验是指施加于一个或多个总体的概率分布或参数的假设. 所作假设可以是正确的, 也可以是错误的.

为判断所作的假设是否正确, 从总体中抽取样本, 根据样本的取值, 按一定原则进行检验, 然后作出接受或拒绝所作假设的决定.

△ 假设检验的内容

- 参数检验（§8.2）
  - 总体均值, 均值差的检验
  - 总体方差, 方差比的检验
- 非参数检验
  - 分布拟合检验（§8.3）
  - 符号检验
  - 秩和检验

△ 假设检验的理论依据

假设检验所以可行, 其理论背景为实际推断原理, 即"小概率原理"

引例1 某产品出厂检验规定：次品率p不超过4%才能出厂．现从一万件产品中任意抽查12件发现3件次品，问该批产品能否出厂？若抽查结果发现1件次品，问能否出厂？

解 假设 $p \le 0.04$,

$p = 0.04$代入

$$P_{12}(3) = C_{12}^{3} p^{3} (1-p)^{9} = 0.0097 < 0.01$$

这是 **小概率事件**，一般在一次试验中是不会发生的，现一次试验竟然发生，故认为原假设不成立，即该批产品次品率$p > 0.04$则该批产品不能出厂．

$$P_{12}(1)=C_{12}^{1}p^{1}(1-p)^{11}=0.306>0.3$$

这不是**小概率事件**,没理由拒绝原假设,从而接受原假设,即该批产品可以出厂.

注1 直接算 $1/12=0.083>0.04$

若不用假设检验,按理不能出厂.

注2 本检验方法是概率意义下的反证法,故拒绝原假设是有说服力的,而接受原假设是没有说服力的.因此应把希望否定的假设作为原假设.

出厂检验问题的数学模型

对总体$X \sim f(x;p)=p^x(1-p)^{1-x},\ x=0,1$提出假设

$$H_0: p \le 0.04; \qquad H_1: p > 0.04$$

要求利用样本观察值

$$(x_1, x_2, \cdots, x_{12}) \qquad (\sum_{i=1}^{12} x_i = 3 \text{ or } 1)$$

对提供的信息作出接受H_0 (可出厂)，还是接受H_1 (不准出厂) 的判断.

引例2　某厂生产的螺钉，按标准强度为 68/mm², 而实际生产的强度 X 服 $N(\mu, 3.6^2)$. 若 $E(X)=\mu=68$, 则认为这批螺钉符合要求，否则认为不符合要求. 为此提出如下假设:

$H_0: \mu = 68$ —— 称为原假设或零假设

原假设的对立面:

$H_1: \mu \neq 68$ —— 称为备择假设

假设检验的任务

必须在原假设与备择假设之间作一选择

现从整批螺钉中取容量为36的样本，其均值为 $\bar{x}=68.5$，问原假设是否正确？

若原假设正确，则 $\overline{X}\sim N(68, 3.6^2/36)$

因而 $E(\overline{X})=68$，即 $\overline{X}$ 偏离68不应该太远，

故 $\left|\dfrac{\overline{X}-68}{3.6/6}\right|$ 取较大值是小概率事件．因此，

可以确定一个常数 c 使得 $P\left(\left|\dfrac{\overline{X}-68}{3.6/6}\right|>c\right)=\alpha$

取 $\alpha=0.05$，则 $c=z_{\frac{\alpha}{2}}=z_{0.025}=1.96$

由 $\left|\dfrac{\bar{X}-68}{3.6/6}\right|>1.96 \Rightarrow \bar{X}>69.18$ 或 $\bar{X}<66.824$

即区间$(-\infty, 66.824)$ 与 $(69.18, +\infty)$ 为检验的**拒绝域**

称 $\bar{X}$ 的取值区间（66.824，69.18）

为检验的**接受域** (实际上没理由拒绝),

现 $\bar{x}=68.5$ 落入接受域, 则接受原假设

$$H_0: \mu = 68$$

由引例2可见，在给定α的前提下，接受还是拒绝原假设完全取决于样本值，因此所作检验可能导致以下两类错误的产生：

第一类错误 —— 弃真错误

第二类错误 —— 取伪错误

假设检验的两类错误

所作判断 / 真实情况	接受 H_0	拒绝 H_0
H_0 为真	正确	第一类错误 (弃真)
H_0 为假	第二类错误 (取伪)	正确

犯第一类错误的概率通常记为 α

犯第二类错误的概率通常记为 β

任何检验方法都不能完全排除犯错误的可能性. 理想的检验方法应使犯两类错误的概率都很小, 但在样本容量给定的情形下, 不可能使两者都很小, 降低一个, 往往使另一个增大.

假设检验的指导思想是控制犯第一类错误的概率不超过α, 然后, 若有必要, 通过增大样本容量的方法来减少 β .

引例2 中，犯第一类错误的概率

$$P(\text{拒绝}H_0 \mid H_0\text{为真})$$

$$= P(\bar{X} < 66.824 \cup \bar{X} > 69.18)$$

$$\boxed{= \alpha = 0.05}$$

若H_0为真，则 $\bar{X} \sim N(68,\ 3.6^2/36)$

所以, 拒绝 H_0 的概率为α， α又称为**显著性水平**, α 越大, 犯第一类错误的概率越大，即越显著.

下面计算犯第二类错误的概率 β

$$\beta = P(\text{接受}H_0 \mid H_0\text{不真})$$

H_0不真, 即$\mu \neq 68$, μ可能小于68, 也可能大于68, β 的大小取决于 μ 的真值的大小.

设 $\mu = 66,\ n = 36,\quad \bar{X} \sim N(66, 3.6^2/36)$

$$\beta_{\mu=66} = P(66.82 \le \bar{X} \le 69.18 \mid \mu = 66)$$

$$= \Phi\left(\frac{69.18-66}{0.6}\right) - \Phi\left(\frac{66.82-66}{0.6}\right)$$

$$= \Phi(5.3) - \Phi(1.37) = 1 - 0.9147 = 0.0853$$

若 $\mu = 69,\ n = 36,\quad \bar{X} \sim N(69, 3.6^2/36)$

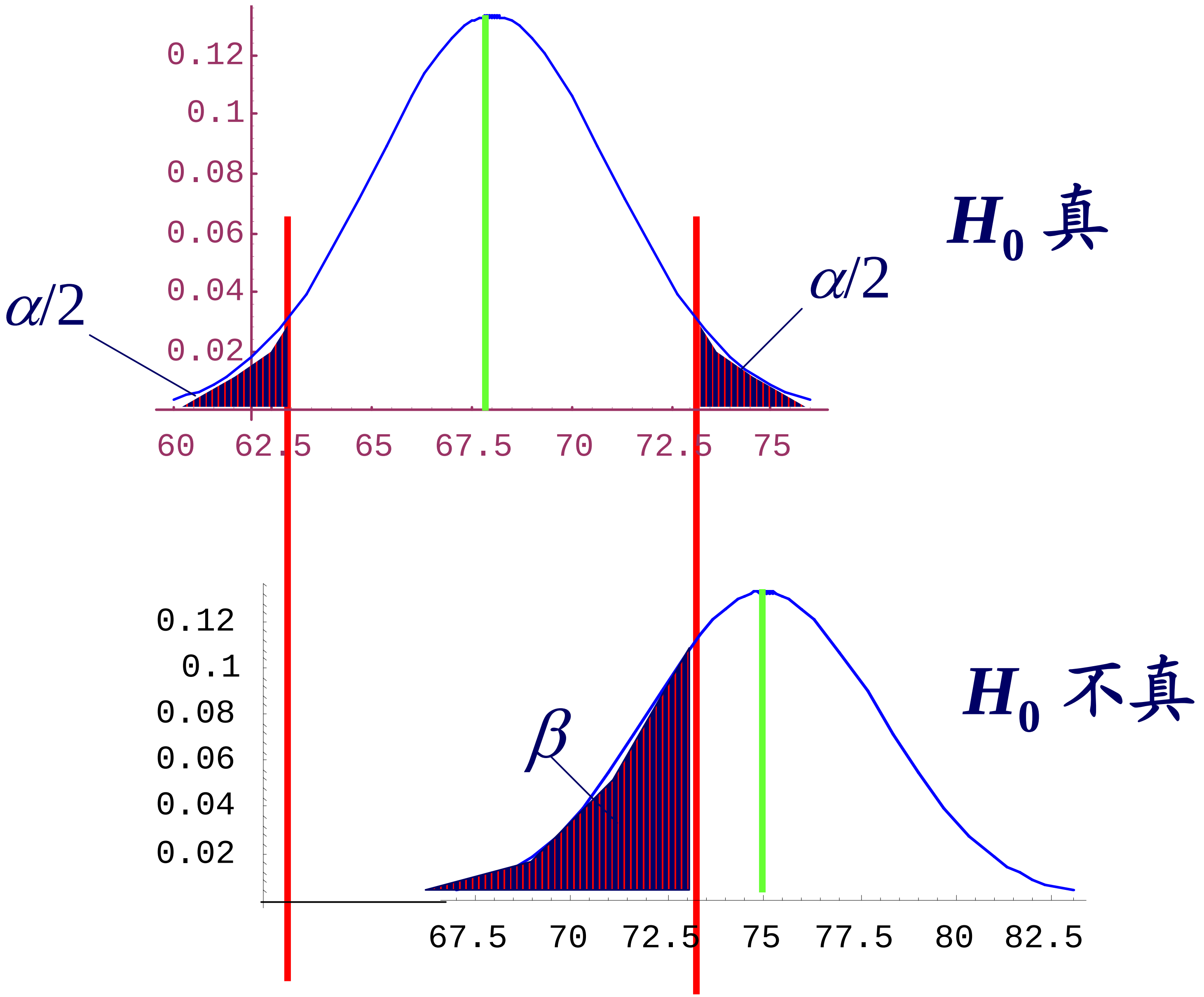
$$\beta_{\mu=69} = P(66.82 \le \bar{X} \le 69.18 \mid \mu = 69)$$

$$= \Phi\left(\frac{69.18 - 69}{0.6}\right) - \Phi\left(\frac{66.82 - 69}{0.6}\right)$$

$$= \Phi(0.3) - \Phi(-3.63)$$

$$= 0.6179 - 0.0002 = 0.6177$$

取伪的概率较大.


0.12
0.1
0.08
0.06
0.04
0.02
60
62.5
65
67.5
70
72.5
75
α/2
α/2
H0 真
0.12
0.1
0.08
0.06
0.04
0.02
67.5
70
72.5
75
77.5
80
82.5
β
H0 不真

现增大样本容量，取$n = 64$, $\mu = 66$, 则

$$\bar{X} \sim N(66, 3.6^2/64)$$

仍取α=0. 05, 则 $c = z_{\frac{\alpha}{2}} = z_{0.025} = 1.96$

由$\left|\dfrac{\bar{X}-68}{3.6/8}\right| > 1.96$ 可以确定拒绝域为

$$(-\infty, 67.118) \text{ 与 } (68.882, +\infty)$$

因此，接受域为(67. 118, 68. 882)

$$\beta_{\mu=66} = P(67.118 \le \bar{X} \le 68.882 \mid \mu = 66)$$

$$\approx \Phi\left(\frac{68.88-66}{0.45}\right) - \Phi\left(\frac{67.12-66}{0.45}\right)$$

$$= \Phi(6.4) - \Phi(2.49)$$

$$\approx 1 - 0.9936 = 0.0064 < 0.0853$$

$$\beta_{\mu=69} = P(67.12 \le \bar{X} \le 68.88 \mid \mu = 69)$$

$$= 0.3936 < 0.6177$$

$$(\mu \to \mu_0, \beta \to 1-\alpha)$$

命题 当样本容量确定后,犯两类错误的概率不可能同时减少.

证 设 $X \sim N(\mu, \sigma_0^2)$ 在水平 α 给定下,检验假设

$$H_0: \mu = \mu_0; \quad H_1: \mu = \mu_1 > \mu_0$$

此时犯第二类错误的概率为

$$\beta = P(\text{接受} H_0 \mid H_0 \text{伪}) = P(\overline{X} - \mu_0 < k \mid \mu = \mu_1)$$

$$= P_{H_1}(\overline{X} - \mu_0 < k) = P_{H_1}(\overline{X} - \mu_1 < k - (\mu_1 - \mu_0))$$

$$= P_{H_1}\left(\frac{\overline{X} - \mu_1}{\sigma_0 / \sqrt{n}} < \frac{k - (\mu_1 - \mu_0)}{\sigma_0 / \sqrt{n}}\right) = \Phi\left(\frac{k - (\mu_1 - \mu_0)}{\sigma_0 / \sqrt{n}}\right)$$

$$\underline{\underline{k=\frac{\sigma_0}{\sqrt{n}}z_\alpha}}\quad \Phi\left(z_\alpha-\frac{\mu_1-\mu_0}{\sigma_0/\sqrt{n}}\right)$$

又 $\beta=\int_{-\infty}^{-z_\beta}\frac{1}{\sqrt{2\pi}}e^{-\frac{x^2}{2}}dx=\Phi(-z_\beta)$ **(见注)**

$\therefore\ z_\alpha-\frac{\mu_1-\mu_0}{\sigma_0/\sqrt{n}}=-z_\beta$ 即 $z_\alpha+z_\beta=\frac{\sqrt{n}}{\sigma_0}(\mu_1-\mu_0)$

由此可见,当 *n* 固定时

1) 若 $\alpha\downarrow\Rightarrow z_\alpha\uparrow\Rightarrow z_\beta\downarrow\Rightarrow\beta\uparrow$

2) 若 $\beta\downarrow\Rightarrow z_\beta\uparrow\Rightarrow z_\alpha\downarrow\Rightarrow\alpha\uparrow$ **证毕.**

注　当 $\mu=\mu_1$ 时 $\overline{X}\sim N(\mu_1,\ \sigma_0^2/n)$

从而 $$\frac{\overline{X}-\mu_1}{\sigma_0/\sqrt{n}}=X^*\sim N(0,\ 1)$$

$$\beta=1-(1-\beta)=1-P(X^*>z_{1-\beta})$$

$$=P(X^*\le z_{1-\beta})=\Phi(z_{1-\beta})=\Phi(-z_\beta)$$

注 1° 一般, 作假设检验时, 先控制犯第一类错误的概率α, 在此基础上使 β 尽量地小. 要降低 β 一般要增大样本容量. 当H_0不真时, 参数值越接近真值, β 越大.

注 2° 备择假设可以是单侧, 也可以双侧.

引例2中的备择假设是双侧的. 若根据以往生产情况, μ_0=68. 现采用了新工艺, 关心的是新工艺能否提高螺钉强度, μ越大越好. 此时可作如下的右边假设检验:

$$H_0: \mu = 68; \qquad H_1: \mu > 68$$

注 3° 关于原假设与备择假设的选取

H_0与H_1地位应平等, 但在控制犯第一类错误的概率 α 的原则下, 使得采取拒绝H_0 的决策变得较慎重, 即H_0 得到特别的保护.

因而, 通常把有把握的、有经验的结论作为原假设, 或者尽可能使后果严重的错误成为第一类错误.

假设检验步骤(三部曲)

- 根据实际问题建立 H_0 与 H_1 .
- 在 H_0 为真时，选择合适统计量 V，由 H_1 确定拒绝域 $\mathfrak{R}$.

$$\mathfrak{R}\begin{cases} \text{双边检验} & (V<V_{1-\frac{\alpha}{2}})\cup(V>V_{\frac{\alpha}{2}}) \\ \text{左边检验} & (V<V_{1-\alpha}) \\ \text{右边检验} & (V>V_{\alpha}) \end{cases}$$

其中
$P(V>V_\alpha)=\alpha$

- 计算，并作出相应判断.

答疑时间

地点中院-312

12月2日 周四　18:00~20:00

12月9日 周四　18:00~20:00

考前答疑

12月27日 13:00~16:00　18:00~20:00

12月28日 13:00~16:00　18:00~20:00