第八章

假设检验

§8.1　假设检验

与单个正态总体参数的假设检验

一、假设检验

1、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．所作的假设可以是正确的，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，按一定的原则进行检验，然后，作出接受或拒绝所作假设的决定．

2、假设检验的内容

- 参数检验
  - 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
  - 符号检验
  - 秩和检验

3、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”

下面通过引例来说明问题

引例1　　某产品的出厂检验规定: 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？

$p=0.04$代入

解　假设 $p \le 0.04$　$P_{12}(3)=C_{12}^{3}p^{3}(1-p)^{9}=0.0097<0.01$

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, 故可认为原假设不成立, 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　直接算 $\frac{1}{12}=0.083>0.04$

若不采用假设检验, 按理也不能够出厂.

上述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

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

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克/mm^2,而实际生产的螺钉强度 X 服从 $N(\mu, 3.6^2)$.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

现从该厂生产的螺钉中抽取容量为 36 的样本,其样本均值为 $\bar{x} = 68.5$,问原假设是否正确?

若原假设正确，则

$$\overline{X} \sim N\left(68, \frac{3.6^2}{36}\right)$$

因而$E(\overline{X}) = 68$,即$\overline{X}$偏离68不应该太远,偏离较远是小概率事件，由于

$$\frac{\overline{X} - 68}{\frac{3.6}{6}} \sim N(0,1)$$

故$\left| \frac{\overline{X} - 68}{\frac{3.6}{6}} \right|$取较大值是小概率事件

规定 α为小概率事件的概率大小,通常取

α= 0.05，0.01,...

因此， 可以确定一个常数 c ， 使得

$$P\left\{\left|\frac{\overline{X}-68}{\frac{3.6}{6}}\right|>c\right\}=\alpha$$

例如， 取 α= 0.05 ， 则

$$c=z_{\frac{\alpha}{2}}=z_{0.025}=1.96$$

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

称 $\overline{X}$ 的取值区间

$$(\ 66.824\ ,\ 69.18\ )$$

为检验的**接受域** (实际上没理由拒绝),而区间

$$(-\infty\ ,66.824\ )\ 与\ (\ 69.18\ ,\ +\infty\ )$$

为检验的**拒绝域**

现 $\overline{x}=68.5$ 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

希望所用的检验方法尽量少犯错误,但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,但在样本的容量给定的情形下，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

$=P\{拒绝H_0 | H_0为真\} = P\{\overline{X} < 66.824 \cup \overline{X} > 69.18\}$

$= \alpha = 0.05$

若H_0为真，则

$$\overline{X} \sim N(68, \frac{3.6^2}{36})$$

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

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

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

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

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

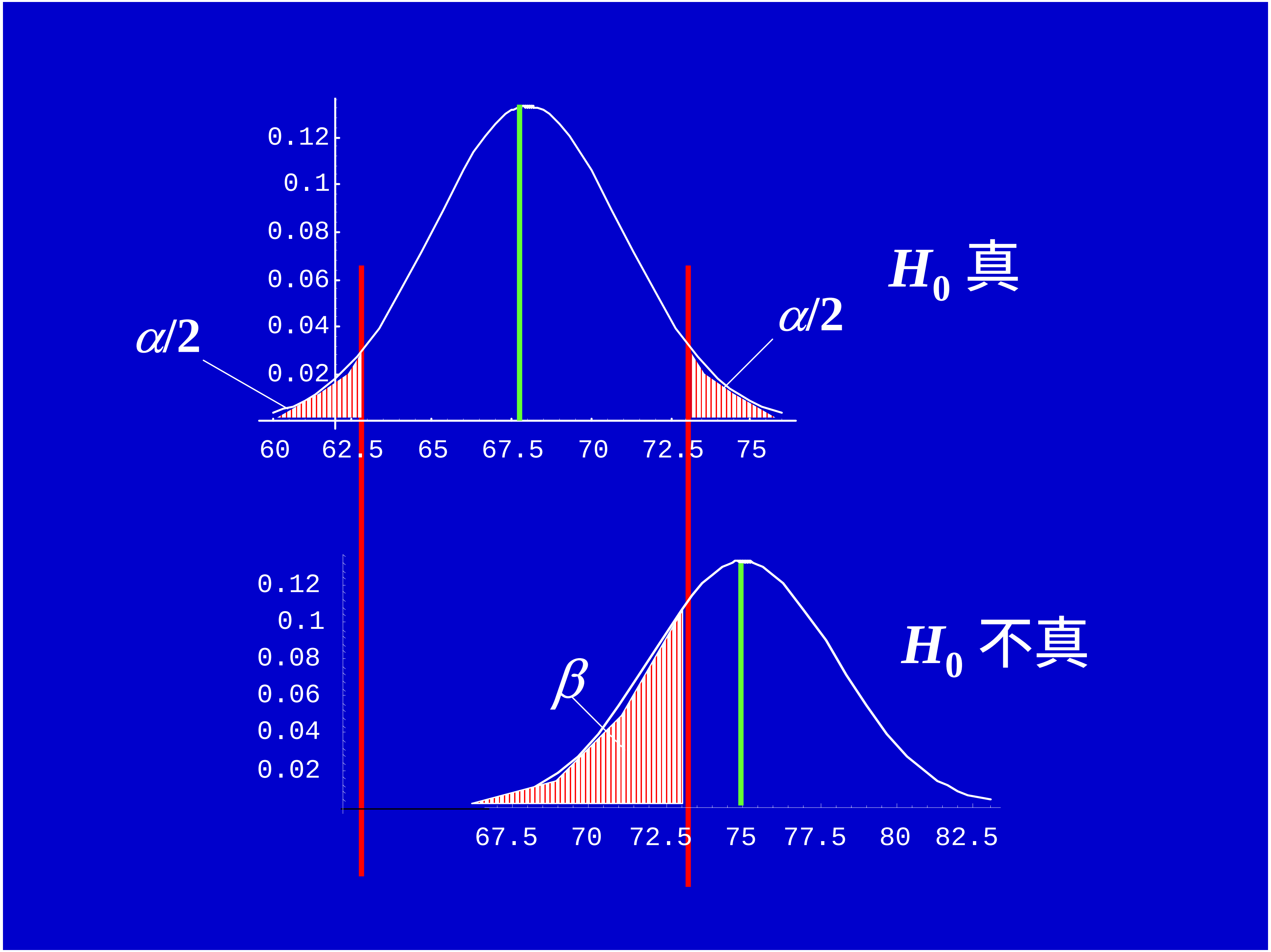
$$\beta_{\mu=69} = P\{66.82 \le \bar{X} \le 69.18 \mid \mu = 69\}$$

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

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

$$= 0.6179 - 0.0002 = 0.6177$$

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

$$\overline{X} \sim N(66, \frac{3.6^2}{64})$$

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

由 $\left| \dfrac{\overline{X} - 68}{\dfrac{3.6}{8}} \right| > 1.96$ 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

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

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

$\therefore z_\alpha-\frac{\mu_1-\mu_0}{\sigma_0/\sqrt{n}}=-z_\beta$ 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 $\beta\downarrow\Rightarrow z_\beta\uparrow\Rightarrow z_\alpha\downarrow\Rightarrow\alpha\uparrow$

注 1° 一般,作假设检验时,先控制犯第一类错误的概率 α, 在保证 α 的条件下使 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,也可以是双侧的.

引例2中的备择假设是双侧的.如果根据以往的生产情况,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, 可作如下的假设检验:

原假设 $H_0: \mu = 68$;备择假设 $H_1: \mu < 68$

当原假设H_0：$\mu = \mu_0 = 68$ 为真时，

$\overline{X} - \mu_0$取较大值的概率较小

当备择假设H_1：$\mu < 68$ 为真时，

$\overline{X} - \mu_0$ 取较大值的概率较大

给定显著性水平α，根据 $P\left\{ \dfrac{\overline{X} - \mu_0}{\sigma / \sqrt{n}} < -z_\alpha \right\} = \alpha$

可确定拒绝域

$$\overline{x} \in \left(-\infty, \mu_0 - z_\alpha \frac{\sigma}{\sqrt{n}}, \quad \right)$$

因而，接受域 $\bar{x} \in (\mu_0 - z_\alpha \dfrac{\sigma}{\sqrt{n}}, +\infty)$

称这种检验为左边检验.

另外，可设 原假设 $H_0: \mu \leq 68$

备择假设 $H_1: \mu > 68$

$$\bar{X} \sim N(\mu, \frac{\sigma^2}{n}), E(\bar{X}) = \mu$$

若原假设正确，则 $P\left\{ \dfrac{\bar{X} - \mu}{\dfrac{\sigma}{\sqrt{n}}} > z_\alpha \right\} = \alpha$

但现不知 μ的真值，只知 $\mu \le \mu_0 = 68$

$$\left\{ \frac{\overline{X} - \mu_0}{\frac{\sigma}{\sqrt{n}}} > z_\alpha \right\} \subset \left\{ \frac{\overline{X} - \mu}{\frac{\sigma}{\sqrt{n}}} > z_\alpha \right\}$$

$$P\left\{ \frac{\overline{X} - \mu_0}{\frac{\sigma}{\sqrt{n}}} > z_\alpha \right\} \le \alpha$$ —— 小概率事件

故取拒绝域 $(\mu_0 + z_\alpha \frac{\sigma}{\sqrt{n}}, +\infty)$ 显著性水平不超过α

注 3°　关于零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)

1、 根据实际问题所关心的内容,建立H_0与H_1

2、 在H_0为真时,选择合适的统计量V,由H_1确定拒绝域形式

给定显著性水平α,其对应的拒绝域

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

其中

$$P\{V > V_{\alpha}\} = \alpha$$

3、 根据样本值计算,并作出相应的判断.

二、单个正态总体参数的假设检验

1、关于μ的检验

拒绝域的推导

给定显著性水平α与样本值$(x_1, x_2, \ldots, x_n)$

设 $X \sim N(\mu, \sigma^2)$，σ^2 已知，需检验：

$$H_0: \mu = \mu_0; \quad H_1: \mu \neq \mu_0$$

构造统计量 $$U = \frac{\bar{X} - \mu_0}{\frac{\sigma}{\sqrt{n}}} \sim N(0,1)$$

$$P\{拒绝H_0 \mid H_0为真\}$$

$$= P\{\,|\overline{X} - \mu_0| \geq k \mid \mu = \mu_0\,\} = P_{H_0}\{\,|\overline{X} - \mu_0| \geq k\,\}$$

$$= P_{H_0}\left\{\left|\frac{\overline{X} - \mu_0}{\sigma/\sqrt{n}}\right| \geq \frac{k}{\sigma/\sqrt{n}}\right\} = P_{H_0}\left\{\left|\frac{\overline{X} - \mu_0}{\sigma/\sqrt{n}}\right| \geq Z_{\frac{\alpha}{2}}\right\} = \alpha$$

$$取 k = Z_{\frac{\alpha}{2}} \frac{\sigma}{\sqrt{n}}$$

所以本检验的拒绝域为

$$_0: \quad |U| \geq Z_{\frac{\alpha}{2}} \quad \text{- - - - - - - - -} \quad \boxed{U\ 检验法}$$

U 检验法 (σ^2 已知)

原假设 H_0	备择假设 H_1	检验统计量及其 H_0为真时的分布	拒绝域
$\mu = \mu_0$	$\mu \neq \mu_0$	$U = \dfrac{\overline{X} - \mu_0}{\dfrac{\sigma}{\sqrt{n}}} \sim N(0,1)$	$\lvert U \rvert > z_{\frac{\alpha}{2}}$
$\mu = \mu_0$	$\mu < \mu_0$		$U < -z_\alpha$
$\mu = \mu_0$	$\mu > \mu_0$		$U > z_\alpha$

T 检验法 (σ^2 未知)

原假设 H_0	备择假设 H_1	检验统计量及其 H_0为真时的分布	拒绝域
$\mu = \mu_0$	$\mu \neq \mu_0$	$T = \dfrac{\overline{X} - \mu_0}{\dfrac{S}{\sqrt{n}}} \sim t(n-1)$	$\lvert T \rvert > t_{\frac{\alpha}{2}}(n-1)$
$\mu = \mu_0$	$\mu < \mu_0$		$T < -t_{\alpha}(n-1)$
$\mu = \mu_0$	$\mu > \mu_0$		$T > t_{\alpha}(n-1)$

例1 某厂生产小型马达，其说明书上写着：这种小型马达在正常负载下平均消耗电流不会超过0.8 安培.

现随机抽取16台马达试验，求得平均消耗电流为0.92安培，消耗电流的标准差为0.32安培.

假设马达所消耗的电流服从正态分布，取显著性水平为 $\alpha = 0.05$，问根据这个样本，能否否定厂方的断言？

解 根据题意待检假设可设为

$H_0: \mu = 0.8$ ； $H_1: \mu > 0.8$

σ未知, 故选检验统计量:

$$T = \frac{\overline{X} - \mu}{S/\sqrt{16}} \sim t(15)$$

查表得 $t_{0.05}(15) = 1.753$, 故拒绝域为

$$\frac{\bar{x} - 0.8}{s/\sqrt{n}} > 1.753 \quad \Rightarrow \bar{x} > 0.8 + 1.753\frac{0.32}{4} = 0.94$$

现 $\bar{x} = 0.92 < 0.94$

故接受原假设, 即不能否定厂方断言.

解二　$H_0: \mu = 0.8$ ；　$H_1: \mu < 0.8$

选用统计量:

$$T = \frac{\overline{X} - \mu}{S/\sqrt{16}} \sim t(15)$$

查表得 $t_{0.05}(15) = 1.753$, 故拒绝域

$$\frac{\bar{x} - 0.8}{s/\sqrt{n}} < -1.753 \Rightarrow \bar{x} < 0.8 - 1.753\frac{0.32}{4} = 0.66$$

现　$\bar{x} = 0.92 > 0.66$

故接受原假设, 即否定厂方断言.

由例1可见：对问题的提法不同(把哪个假设作为原假设)，统计检验的结果也会不同.

由于假设检验是控制犯第一类错误的概率，使得拒绝原假设 H_0 的决策变得比较慎重，也就是 H_0 得到特别的保护. 因而，通常把有把握的，经验的结论作为原假设，或者尽量使后果严重的错误成为第一类错误.

上述两种解法的立场不同，因此得到不同的结论.第一种假设是不轻易否定厂方的结论；第二种假设是不轻易相信厂方的结论.

2、关于σ^2的检验　　χ^2检验法

原假设 H_0	备择假设 H_1	检验统计量及其在 H_0为真时的分布	拒绝域
$\sigma^2=\sigma_0^2$	$\sigma^2\neq\sigma_0^2$	$\chi^2=\dfrac{\sum_{i=1}^{n}(X_i-\mu)^2}{\sigma_0^2}\sim\chi^2(n)$ (μ 已知)	$\chi^2\le\chi^2_{1-\frac{\alpha}{2}}(n)$ 或 $\chi^2\ge\chi^2_{\frac{\alpha}{2}}(n)$
$\sigma^2=\sigma_0^2$	$\sigma^2<\sigma_0^2$		$\chi^2\le\chi^2_{1-\alpha}(n)$
$\sigma^2=\sigma_0^2$	$\sigma^2>\sigma_0^2$		$\chi^2\ge\chi^2_{\alpha}(n)$

原假设 H_0	备择假设 H_1	检验统计量及其在 H_0为真时的分布	拒绝域
$\sigma^2=\sigma_0^2$	$\sigma^2\neq\sigma_0^2$	$\chi^2=\frac{(n-1)S^2}{\sigma_0^2}\sim\chi^2(n-1)$ （μ 未知）	$\chi^2<\chi^2_{1-\frac{\alpha}{2}}(n-1)$ 或 $\chi^2>\chi^2_{\frac{\alpha}{2}}(n-1)$
$\sigma^2=\sigma_0^2$	$\sigma^2<\sigma_0^2$		$\chi^2<\chi^2_{1-\alpha}(n-1)$
$\sigma^2=\sigma_0^2$	$\sigma^2>\sigma_0^2$		$\chi^2>\chi^2_{\alpha}(n-1)$

例2　　某汽车配件厂在新工艺下对加工好的25个活塞的直径进行测量,得样本方差S^2=0.00066. 已知老工艺生产的活塞直径的方差为0.00040. 问进一步改革的方向应如何？

解　　一般进行工艺改革时, 若指标的方差显著增大, 则改革需朝相反方向进行以减少方差；若方差变化不显著, 则需试行别的改革方案.

设测量值 $X \sim N(\mu, \sigma^2)$, $\sigma^2 = 0.00040$

需考察改革后活塞直径的方差是否不大于改革前的方差？故待检验假设可设为：

$$H_0: \sigma^2 = 0.00040 \; ; \; H_1: \sigma^2 > 0.00040.$$

此时可采用效果相同的单边假设检验

$$H_0: \sigma^2 = 0.00040 \; ; \; H_1: \sigma^2 > 0.00040.$$

取统计量 $\chi^2 = \dfrac{(n-1)S^2}{\sigma_0^2} \sim \chi^2(n-1)$

拒绝域 $\Re_0$： $\chi^2 > \chi_\alpha^2(n-1) = \chi_{0.05}^2(24) = 36.415$

$$\chi_0^2 = \frac{24 \times 0.00066}{0.00040} = 39.6 > 36.415$$

落在$\Re_0$内, 故拒绝H_0. 即改革后的方差显著大于改革前的方差, 因此下一步的改革应朝相反方向进行.