第8章 假设检验

第一节　　假设检验

- 假设检验的基本思想和方法
- 假设检验的一般步骤
- 假设检验的两类错误
- 课堂练习
- 小结 布置作业

一、假设检验的基本思想和方法

在本节中，我们将讨论不同于参数估计的另一类重要的统计推断问题. 这就是根据样本的信息检验关于总体的某个假设是否正确.

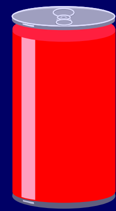
这类问题称作假设检验问题 .

假设检验 $\begin{cases} \text{参数假设检验} \\ \text{非参数假设检验} \end{cases}$

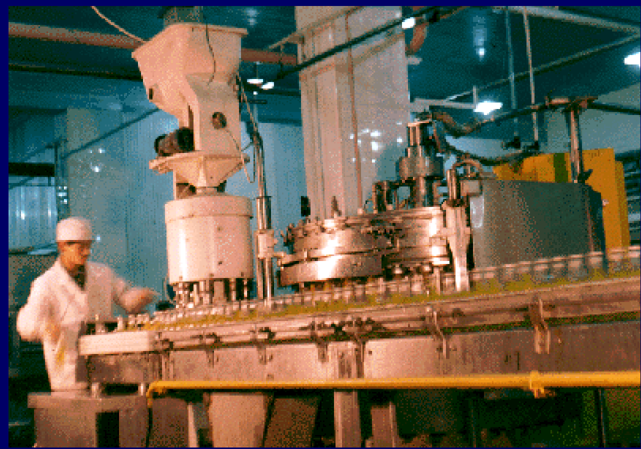
总体分布已知，检验关于未知参数的某个假设

总体分布未知时的假设检验问题

罐装可乐的容量按标准应在350毫升和360毫升之间.



生产流水线上罐装可乐不断地封装，然后装箱外运. 怎么知道这批罐装可乐的容量是否合格呢？



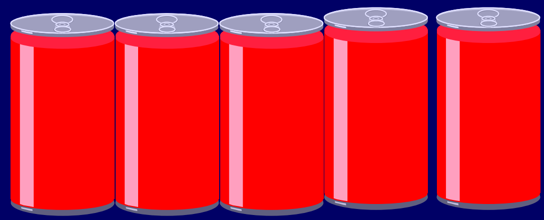
把每一罐都打开倒入量杯，看看容量是否合于标准.

这样做显然不行！

通常的办法是进行抽样检查.

每隔一定时间，抽查若干罐 . 如每隔1小时，抽查5罐，得5个容量的值X_1，…，X_5，根据这些值来判断生产是否正常.

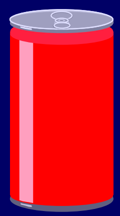
如发现不正常，就应停产，找出原因，排除故障，然后再生产；如没有问题，就继续按规定时间再抽样，以此监督生产，保证质量.



很明显，不能由5罐容量的数据，在把握不大的情况下就判断生产 不正常，因为停产的损失是很大的.

当然也不能总认为正常，有了问题不能及时发现，这也要造成损失.

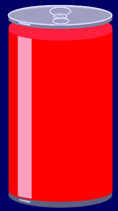
如何处理这两者的关系，假设检验面对的就是这种矛盾.



罐装可乐的容量按标准应在350毫升和360毫升之间.

现在我们就来讨论这个问题.

在正常生产条件下，由于种种随机因素的影响，每罐可乐的容量应在355毫升上下波动. 这些因素中没有哪一个占有特殊重要的地位. 因此，根据中心极限定理，假定每罐容量服从正态分布是合理的.



这样，我们可以认为$X_1,\ldots,X_5$是取自正态总体 $N(\mu,\sigma^2)$ 的样本，当生产比较稳定时，σ^2 是一个常数. 现在要检验的假设是：

H_0: $\mu = \mu_0$ （$\mu_0 = 355$）

在实际工作中，往往把不轻易否定的命题作为原假设.

它的对立假设是：

H_1: $\mu \neq \mu_0$

注意：原假设总是有等号：=或≤ 或 ≥。

称H_0为原假设（或零假设，解消假设）；

称H_1为备选假设（或对立假设）.

那么，如何判断原假设H_0是否成立呢？

由于μ是正态分布的期望值，它的估计量是样本均值$\overline{X}$，因此可以根据$\overline{X}$与μ_0的差距$\left|\overline{X}-\mu_0\right|$来判断$H_0$是否成立.

当$|\overline{X}-\mu_0|$较小时，可以认为H_0是成立的；

当$|\overline{X}-\mu_0|$较大时，应认为H_0不成立，即生产已不正常.

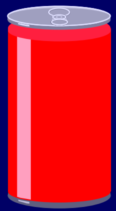
较大、较小是一个相对的概念，合理的界限在何处？应由什么原则来确定？

问题归结为对差异作定量的分析，以确定其性质.

差异可能是由抽样的随机性引起的，称为

“抽样误差”或 随机误差

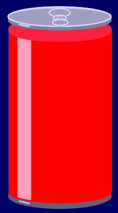
这种误差反映偶然、非本质的因素所引起的随机波动.



然而，这种随机性的波动是有一定限度的，如果差异超过了这个限度，则我们就不能用抽样的随机性来解释了.

必须认为这个差异反映了事物的本质差别，即反映了生产已不正常.

这种差异称作 “系统误差”



问题是，根据所观察到的差异，如何判断它究竟是由于偶然性在起作用，还是生产确实不正常？

即差异是“抽样误差”还是“系统误差”所引起的？

这里需要给出一个量的界限．

假设检验所依据的基本原理是小概率原理。
问题是：如何给出这个量的界限？

这里用到人们在实践中普遍采用的一个原则：

小概率事件在一次试验中基本上不会发生．

假设检验所依据的基本原理是小概率原理。

在假设检验中，我们称这个小概率为显著性水平，用α表示.

α的选择要根据实际情况而定。

常取 $\alpha = 0.1, \alpha = 0.01, \alpha = 0.05.$

现在回到我们前面罐装可乐的例中：

在提出原假设H_0后，如何作出接受和拒绝H_0的结论呢？

罐装可乐的容量按标准应在350毫升和360毫升之间．一批可乐出厂前应进行抽样检查，现抽查了n 罐，测得容量为 $X_1,X_2,\ldots,X_n$，问这一批可乐的容量是否合格？

提出假设

$$H_0:\ \mu = 355 \longleftrightarrow H_1:\ \mu \neq 355$$

由于σ已知，

选检验统计量 $U = \dfrac{\overline{X} - \mu_0}{\sigma/\sqrt{n}} \sim N(0,1)$

它能衡量差异 $|\overline{X} - \mu_0|$ 大小且分布已知.

对给定的显著性水平，α 可以在$N(0,1)$表中查到分位点的值 $u_{\alpha/2}$，使

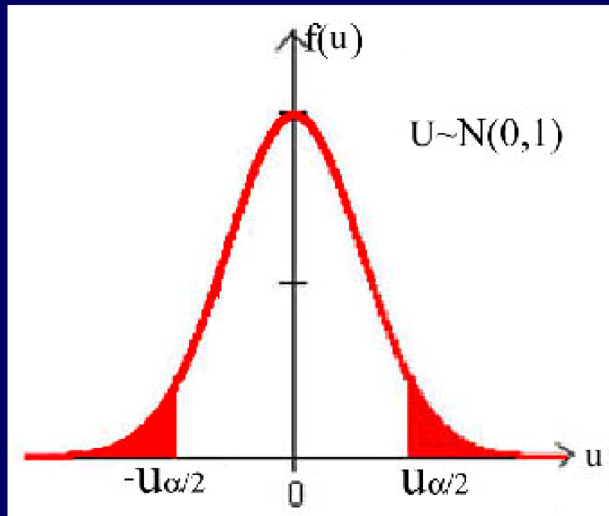
$$P\{|U| > u_{\alpha/2}\} = \alpha$$

$$P\{|U|>u_{\alpha/2}\}=\alpha$$

也就是说,“ $|U|>u_{\alpha/2}$ ”是一个小概率事件.

故我们可以取拒绝域为:

$$W: \ |U|>u_{\alpha/2}$$



如果由样本值算得该统计量的实测值落入区域 W，则拒绝 H_0；否则，不能拒绝 H_0.

这里所依据的逻辑是：

如果H_0是对的，那么衡量差异大小的某个统计量落入区域 W(拒绝域) 是个小概率事件. 如果该统计量的实测值落入W，也就是说，H_0成立下的小概率事件发生了，那么就认为H_0不可信而否定它. 否则我们就不能否定H_0（只好接受它）.

不否定H_0并不是肯定H_0一定对，而只是说差异还不够显著，还没有达到足以否定H_0的程度．

所以假设检验又叫

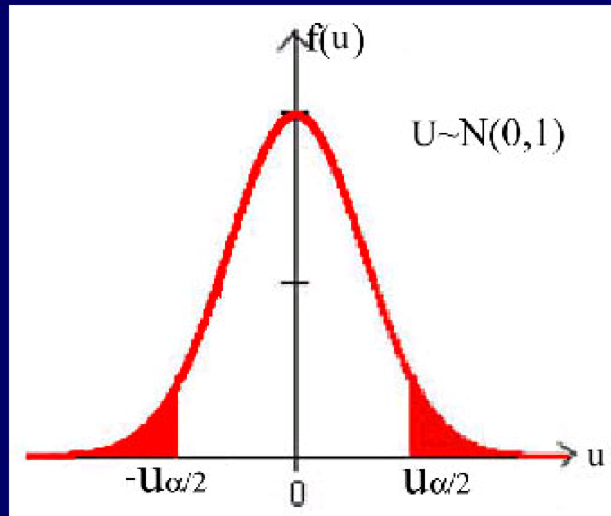
“显著性检验”

如果显著性水平 α 取得很小，则拒绝域也会比较小.

其产生的后果是：

H_0难于被拒绝.

如果在 α 很小的情况下H_0仍被拒绝了，则说明实际情况很可能与之有显著差异.



基于这个理由，人们常把$\alpha=0.05$时拒绝H_0称为是显著的，而把在 $\alpha=0.01$时拒绝H_0称为是高度显著的.

4、假设检验中的拒绝域和接受域

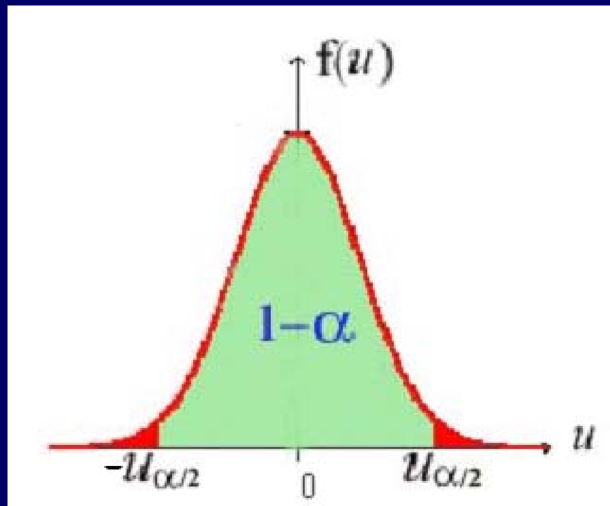
- 在规定了检验的显著性水平α后，根据容量为n的样本，按照统计量的理论概率分布规律，可以确定据以判断拒绝和接受原假设的检验统计量的临界值。
- 临界值将统计量的所有可能取值区间分为两个互不相交的部分，即原假设的拒绝域和接受域。

单、双边检验

H_0：$\mu=\mu_0$，H_1：$\mu \neq \mu_0$ ，拒绝域取在两侧，称为**双边检验**。

H_0：$\mu \leq \mu_0$，H_1：$\mu > \mu_0$ ，拒绝域取在右侧，称为**右边检验**。

H_0：$\mu \geq \mu_0$，H_1：$\mu < \mu_0$ ，拒绝域取在左侧，称为**左边检验**。



前面可乐的例子就是双边检验。
下面看一个单侧检验的例子.

二、假设检验的一般步骤

1. 根据实际问题的需要，提出原假设H_0和备择假设H_1
2. 给定显著性水平α以及样本容量n
3. 确定检验统计量以及拒绝域的形式
4. 按$P\{当H_0为真拒绝H_0\} \leq \alpha$求出拒绝域
5. 取样，根据样本观察值做出决策，是拒绝H_0，还是接受H_0

检验名称

一般说来，按照检验所用的统计量的分布，分为

Z检验/ U 检验	用正态分布
t 检验	用 t 分布
χ^2 检验	用 χ^2 分布
F 检验	用 F 分布

- 例如，工厂生产的某产品次品率不超过5%才能出厂。今抽检100件产品，发现次品4件，问这批产品能否出厂？要求检验结果具有95%的置信度。

解 令

$$X=\begin{cases}0, & \text{检验合格}\\ 1, & \text{检验不合格}\end{cases}$$

则总体 $X \sim b(1,p)$. $E(X)=p$, $\therefore$ p 的 $1-\alpha$ 的置信区间为

$$\left(\bar{X}-\frac{S}{\sqrt{n}}z_{\alpha/2},\ \bar{X}+\frac{S}{\sqrt{n}}z_{\alpha/2}\right)$$

现 $n=100$, $\bar{X}=0.04$, $z_{\alpha/2}=z_{0.025}=1.96$，又因 $X^2=X$, 故

$$S^2=\frac{1}{n-1}\left(\sum_{i=1}^{n}X_i^2-n\bar{X}^2\right)=\frac{n}{n-1}(\bar{X}-\bar{X}^2)=\frac{n}{n-1}\bar{X}(1-\bar{X})=0.0388$$

求得 p 的 95% 置信区间为（0.0014,0.0782）

由于置信上限 0.0782 > 5%，故这批产品不能出厂

例 甲从乙厂定购了一批产品,双方约定产品次品率**p**不能超过5%,否则甲方有权拒收. 验收时随机抽检了100件产品,发现次品6件,问甲方能否接收该批产品? $(\alpha=0.05)$

解 令

$$X=\begin{cases}0\text{, 检验合格}\\1\text{, 检验不合格}\end{cases}$$

则总体 $X\sim b(1,p)$,按题意,要检验假设

$$H_0: p\le p_0=0.05\ ,\ H_1: p>p_0$$

当H_0成立时,由大样本理论有 $\dfrac{\bar{X}-p_0}{S/\sqrt{n}}\overset{\text{近似}}{\sim}N(0,1)$

采用 u 检验法,求得 H_0 的拒绝域为 $\bar{X}>p_0+\dfrac{S}{\sqrt{n}}Z_\alpha$

计算得 $s=\sqrt{\dfrac{\text{n}}{\text{n-1}}\bar{x}(1-\bar{x})}=0.2387,\ p_0+\dfrac{s}{\sqrt{n}}Z_{0.05}=0.089$

$\because\ \bar{x}=0.06<0.089\ \therefore$不拒绝$H_0$,即甲方可以接收该批产品.

假设检验和区间估计的关系（8.4）

【同】均是用于总体参数推断的统计方法；应用的抽样分布原理相同（借助的统计量和分布相同）；

【异】

1. 目的不同：区间估计解决的是定量问题，假设检验解决的是定性问题；
2. 使用方法不同：区间估计使用“顺推法”，假设检验使用“反证法”；
3. 置信区间不同：区间估计是以 $\overline{X}$ 为基准构造的随机区间；假设检验是以$\mu=\mu_0$为基准构造的常数区间；
4. 适用场合不同：区间估计之前不需要了解未知参数的有关信息；假设检验则应对未知参数有所了解，只是对作出的某种推断无确切的把握。

例 甲从乙厂定购了一批产品, 双方约定产品次品率p不能超过5%,否则甲方有权拒收. 验收时随机抽检了100件产品, 发现次品4件,问甲方能否接收该批产品? $(\alpha = 0.05)$

改为4件

解 令 $X = \begin{cases} 0, & \text{检验合格} \\ 1, & \text{检验不合格} \end{cases}$

则总体 $X \sim b(1, p)$,按题意,要检验假设

$$H_0: p \geq p_0 = 0.05,\ H_1: p < p_0$$

当H_0成立时, 由大样本理论有 $\dfrac{\bar{X} - p_0}{S/\sqrt{n}} \overset{\text{近似}}{\sim} N(0,1)$

采用 u 检验法, 求得H_0的拒绝域为 $\bar{X} < p_0 - \dfrac{S}{\sqrt{n}} Z_\alpha$

计算得 $s = \sqrt{\dfrac{\text{n}}{\text{n-1}} \bar{x}(1-\bar{x})} = 0.1969,\ p_0 - \dfrac{s}{\sqrt{n}} Z_{0.05} = 0.0176$

$\because\ \bar{x} = 0.04 > 0.0176 \therefore$不拒绝$H_0$,即甲方可不接收该批产品.

问题: 甲从乙厂定购了一批产品,双方约定产品次品率p不能超过5% , 否则甲方有权拒收.

① 随机抽检100件产品, 发现次品6件,若检验假设

$$H_0: p \le p_0 = 0.05 ,\ H_1: p > p_0$$

结论 不拒绝H_0, 甲方可以接收该批产品

② 随机抽检100件产品,发现次品4件,若检验假设

$$H_0: p \ge p_0 = 0.05 ,\ H_1: p < p_0$$

结论 不拒绝H_0,甲方不接收该批产品

问

1. 为什么两个结论不一致?
2. 哪个结论可信度高?

问题： 甲从乙厂定购了一批产品,双方约定产品次品率p不能超过5%，否则甲方有权拒收.

① 随机抽检100件产品, 发现次品6件,若检验假设

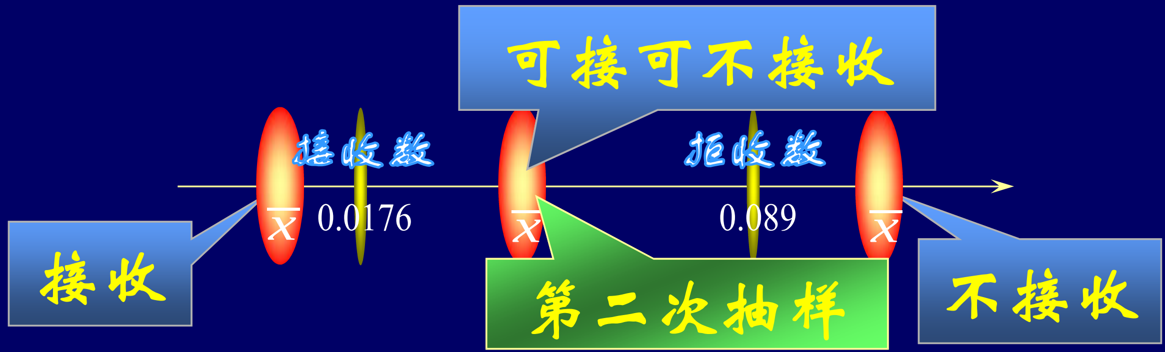
$$H_0: p \le p_0 = 0.05,\ H_1: p > p_0$$

结论 不拒绝H_0, 甲方可以接收该批产品

② 随机抽检100件产品,发现次品4件,若检验假设

$$H_0: p \ge p_0 = 0.05,\ H_1: p < p_0$$

结论 不拒绝H_0,甲方不接收该批产品



《计数抽样检验程序》GB/T2828.1-2003/ISO2859-1:1999

三、假设检验的两类错误

假设检验会不会犯错误呢？

由于作出结论的依据是下述

小概率原理

不是一定不发生

小概率事件在一次试验中基本上不会发生．

如果H_0成立，但统计量的实测值落入否定域，从而作出否定H_0的结论，那就犯了“以真为假”的错误.

如果H_0不成立，但统计量的实测值未落入否定域，从而没有作出否定H_0的结论，即接受了错误的H_0，那就犯了“以假为真”的错误.

请看下表

假设检验的两类错误

	实际情况	
决定	H_0为真	H_0不真
拒绝H_0	第一类错误	正确
接受H_0	正确	第二类错误

犯两类错误的概率:

$$P\{\text{拒绝}H_0|H_0\text{为真}\}=\alpha,$$

$$P\{\text{接受}H_0|H_0\text{不真}\}=\beta.$$

显著性水平α为犯第一类错误的概率.

假设检验中的两类错误

（决策结果）

H_0: 无罪

假设检验就好像一场审判过程

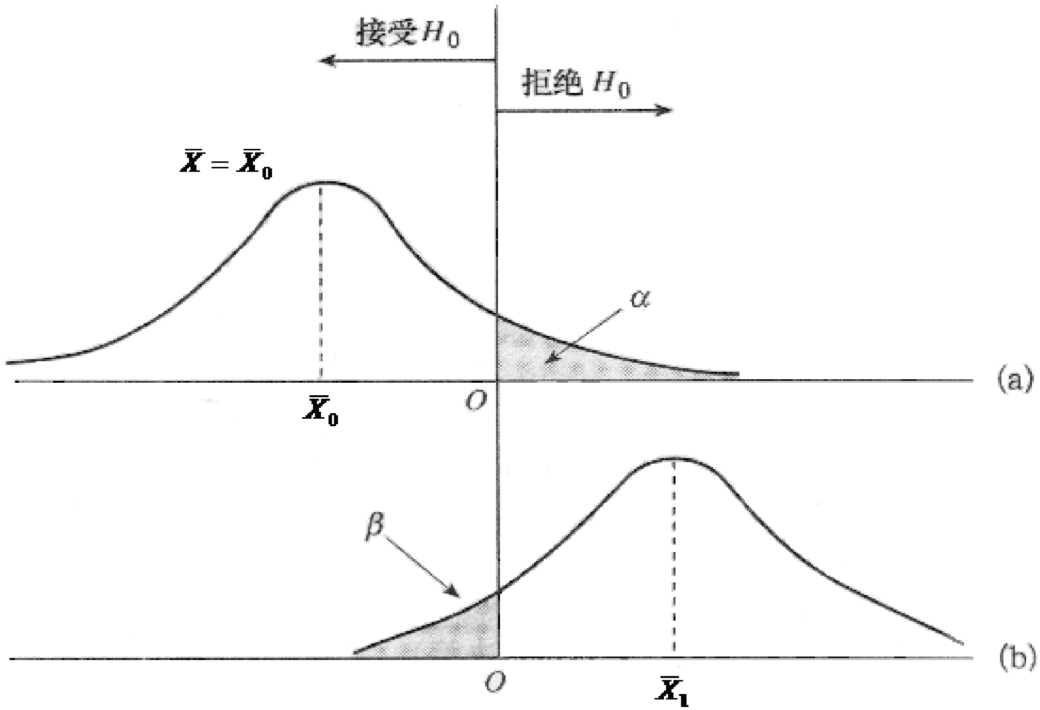
统计检验过程

陪审团审判		
裁决	实际情况	
	无罪	有罪
无罪	正确	错误
有罪	错误	正确

H_0 检验		
决策	实际情况	
	H_0为真	H_0为假
接受H_0	正确决策 $(1-\alpha)$	第二类错误(β)
拒绝H_0	第一类错误(α)	正确决策 $(1-\beta)$

假设检验两类错误关系示意图

以单侧上限检验为例，设H_0： $\bar{X} \leq \bar{X}_0$，H_1： $\bar{X} > \bar{X}_0$



图(a)
H_0为真时，
$\bar{X}_i$的分布

图(b)
H_1为真时，
$\bar{X}_i$的分布

从上图可以看出，如果临界值沿水平方向右移，**α**将变小而**β**变大，即若减小**α**错误，就会增大犯**β**错误的机会；如果临界值沿水平方向左移，**α**将变大而**β**变小，即若减小**β**错误，也会增大犯**α**错误的机会。

两类错误的概率的关系

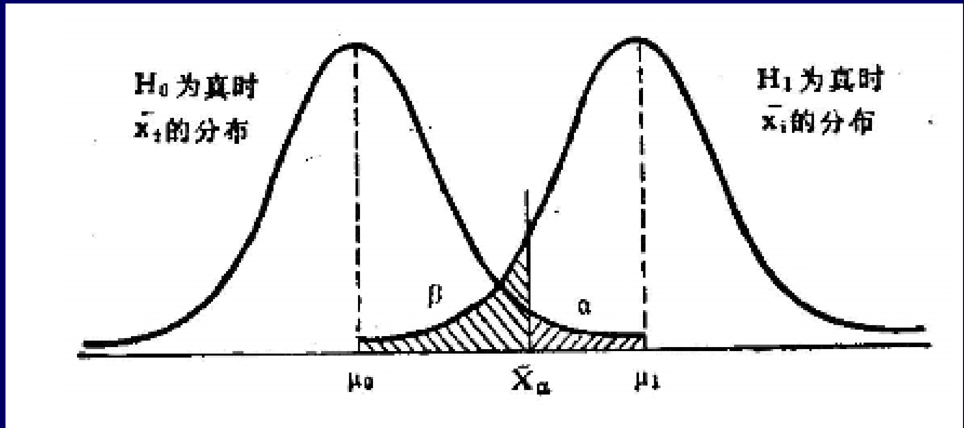
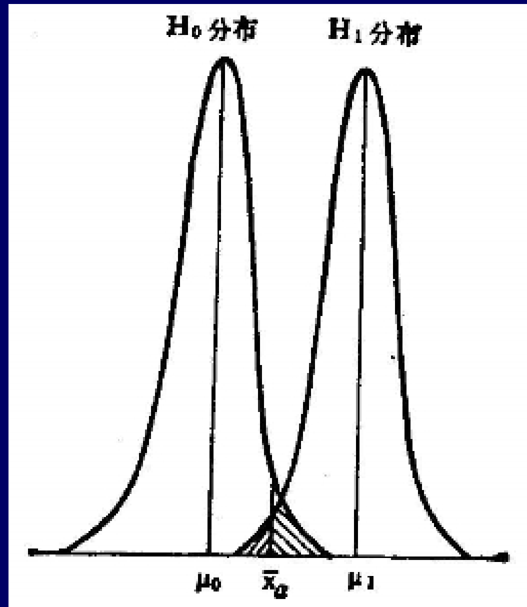


图7-2 α与β的关系示意图



样本均值分布的标准误差为 $\sigma/\sqrt{n}$

两类错误是互相关联的， 当样本容量固定时，一类错误概率的减少导致另一类错误概率的增加.

要同时降低两类错误的概率 α, β 或者要在 α 不变的条件下降低 β，需要增加样本容量.

四、小结

提出假设

根据统计调查的目的，提出原假设H_0和备选假设H_1

抽取样本

检验假设

对差异进行定量的分析，确定其性质(是随机误差还是系统误差．为给出两者界限，找一检验统计量T，在H_0成立下其分布已知.）

显著性水平α

$P(T \in W) = \alpha$

α ------犯第一类错误的概率，W为拒绝域

作出决策

拒绝还是不能拒绝H_0