第八章　假设检验

假设检验问题: 根据样本的信息检验关于总体的某个假设是否正确。

假设检验问题分类:

假设检验
- 参数假设检验 —— 总体分布已知，检验关于未知参数的某个假设
- 非参数假设检验 —— 总体分布未知时的假设检验问题

第一节　假设检验

一．假设检验的基本思想

设总体 X 含有未知参数θ (或总体分布函数 F(x) 未知)检验下述假设:

假设 H_0： $\theta = \theta_0$ 或 $F(x) = F_0(x)$ 其中：

θ_0 是某个已知常数 (或 $F_0(x)$ 是某个已知的分布函数)。则抽取容量为 n 的样本，利用样本提供的信息对假设 H_0 作出判断，从而确定是否接受 H_0

例如： $X \sim N(\mu,\sigma^2)$， μ 未知

检验假设： H_0： $\mu = \bar{x}$

根据上一章的讨论，显然 H_0 是可以被接受的，因为 $\bar{X}$ 是总体 X 的待估计参数 μ 的无偏估计。

二. 判断“假设”的根据

不是一定不发生

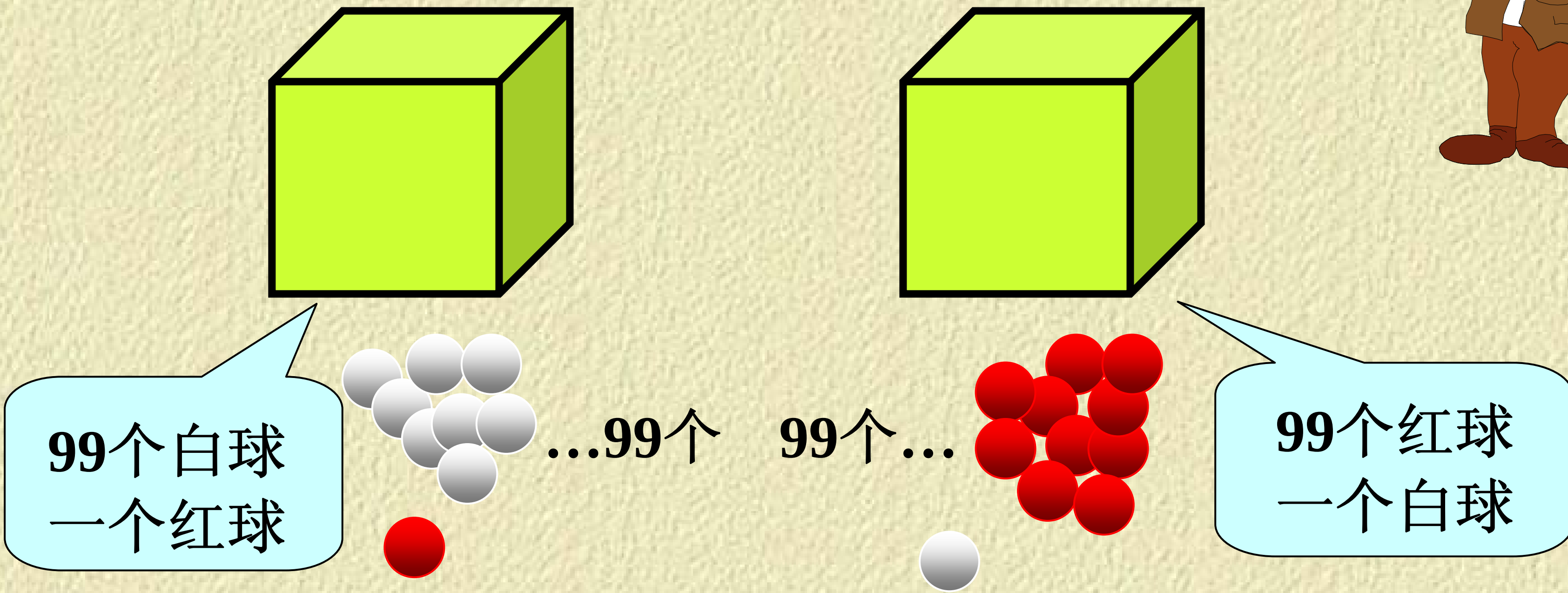
—— 小概率事件原理

小概率事件在一次试验中是几乎不可能发生的

如果在假设 H_0 成立的条件下某事件是小概率事件，但在一次试验中却发生了，于是就可怀疑假设 H_0 的正确性从而拒绝 H_0

现用一个例子来说明这个原则.

例如：现有两个盒子，各装有100个球.

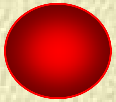


现从两盒中随机取出一个盒子

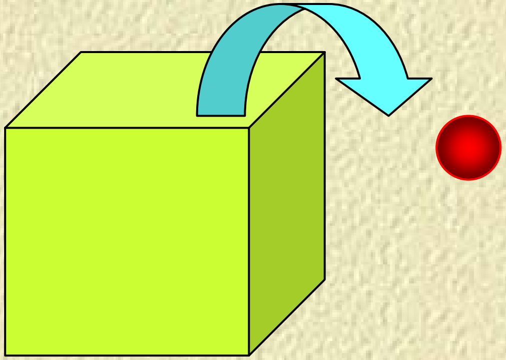
问：这个盒子里是白球 99个还是红球 99 个？

若假设：这个盒子里有 99 个白球.

当从中随机摸出一个球时，发现是红球：



此时应如何判断这个假设是否成立呢？



假设其中真有 99 个白球，摸出红球的概率只有 1/100，这是小概率事件.

但此小概率事件在一次试验中竟然发生了，这就不得不怀疑所作的假设。

注：这个例子中所使用的推理方法，可以称为是带概率性质的反证法。但它不同于一般的反证法。

一般的反证法要求在原假设成立的条件下导出的结论是绝对成立的，如果事实与之矛盾，则完全绝对地否定原假设。

概率反证法的逻辑是：如果小概率事件在一次试验中居然发生了，则就可以以很大的把握否定原假设，否则就不能否定原假设。

在假设检验中，常称这个小概率为显著性水平，用 α 表示。

三. 假设检验的两类错误

1. 第一类错误 (弃真): 如果 H_0 是正确的，但却被错误地否定了。

2. 第二类错误 (取伪): 如果 H_0 是不正确的，但却被错误地接受了。

若设 犯两类错误的概率分别为:

$$P\{\text{拒绝}H_0 \mid H_0\text{为真}\} = \alpha$$

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

则显著性水平 α 为犯第一类错误的概率。

上页 下页 返回

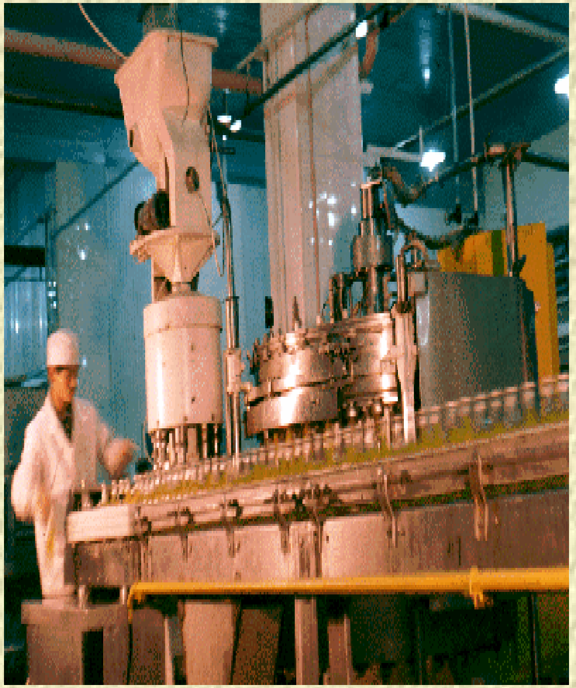
注：两类错误是互相关联的，当样本容量 n 固定时，一类错误概率的减少必导致另一类错误概率的增加。

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 n

在实际问题中，通常的做法是：

先对犯第一类错误（弃真）的概率加以控制，同时再考虑使犯第二类错误（取伪）的概率尽可能的小。

四. 假设检验的具体做法



例1. 罐装可乐容量的检验问题

在一条生产可乐的流水线上罐装可乐不断地封装，然后装箱外运。罐装可乐的容量按标准应在 **350** 毫升和 **360** 毫升之间。

试问：如何检验这批罐装可乐的容量是否合格呢？

分析：若把每一罐可乐都打开倒入量杯，看看容量是否合于标准。这显然是不可行的。

上页 下页 返回

通常的办法是：进行抽样检查.

即，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.

如发现不正常则应停产，找出原因，排除故障，然后再生产；如生产正常，就继续按规定时间再抽样，以此监督生产，保证质量。

显然：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？

现用假设检验的方法来处理这对矛盾

注意到：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.

故：可以认为样本是取自正态总体 $N(\mu, \sigma^2)$

现抽查了n 罐，测得容量为：$X_1, X_2, \cdots, X_n$

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: $\mu = \mu_0$ ($\mu_0 = 355$)

它的对立假设是:

H_1: $\mu \neq \mu_0$

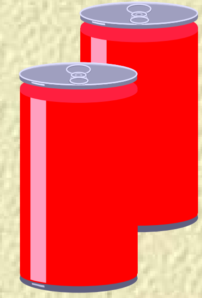
在实际问题中，往往把不轻易否定的命题作为原假设.

称 H_0 为原假设（或零假设

称 H_1 为备择假设（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，它的无偏估计量是样本均值 $\bar{X}$，因此可以根据 $\bar{X}$ 与 μ_0 的差距 $|\bar{X} - \mu_0|$ 来判断 H_0 是否成立.



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

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

而较大、较小是一个相对的概念，那么它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.

注意到：当差异是由抽样的随机性引起时，则称其为“抽样误差”或随机误差；它反映了由偶然、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

上页 下页 返回

如果差异超过了这个限度，则就不能用抽样的随机性来解释了。此时可认为这个差异反映了事物的本质差别，则称其为“系统误差”

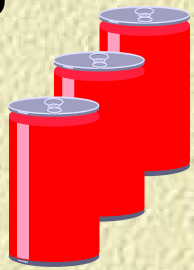
从而问题就转化为：

如何判断差异是由“抽样误差”还是“系统误差”所引起的？

解决的方法：给出一个量的界限，即显著性水平α

从而提出假设：

$$H_0: \mu = \mu_0 = 355 \quad \longleftrightarrow \quad H_1: \mu \neq \mu_0 = 355$$

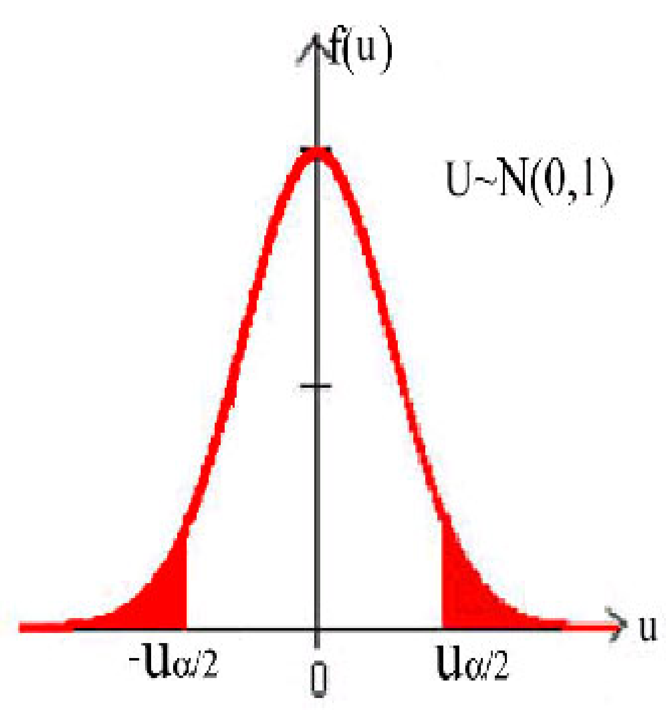


因为 σ 已知，所以构造统计量为：

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

对给定的显著性水平α，查正态分布的上分位α点的值 $u_{\alpha/2}$，使：

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



即 $|U|>u_{\alpha/2}$ 是一个小概率事件

故可以取拒绝域 C 为：

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

如果由样本值算得该统计量的实测值落入区域 C，则拒绝 H_0；否则就接受 H_0.

注：▲ 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 C (拒绝域) 是个小概率事件。

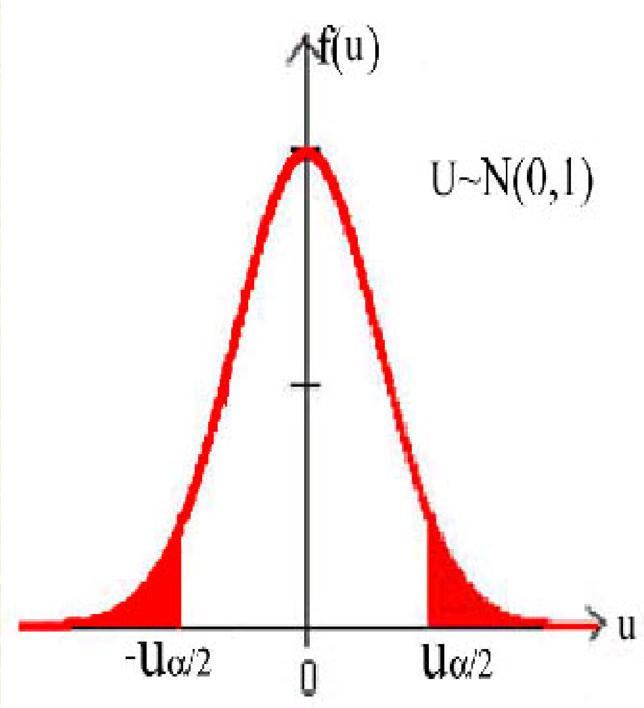
如果该统计量的实测值落入 C，即 H_0 成立下的小概率事件发生了，那么就认为 H_0 不可信而否定它；否则就不能否定 H_0 而只好接受 H_0

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 。

故假设检验又称为“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。



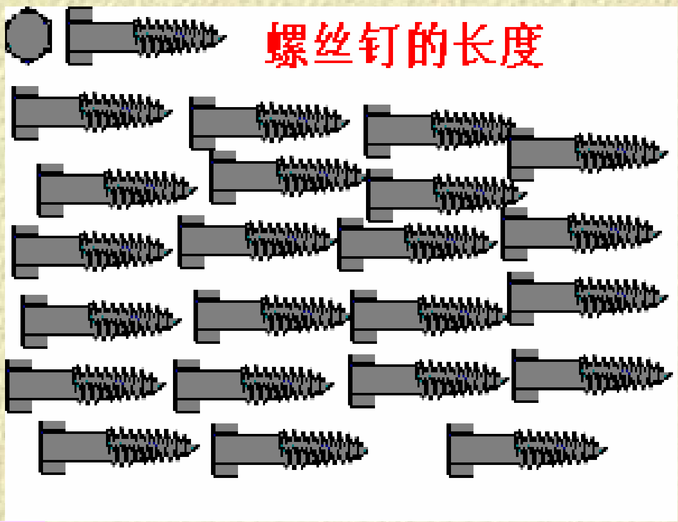
基于这个理由，人们常把$\alpha = 0.05$ 时拒绝 H_0称为是**显著的**。

把在$\alpha = 0.01$ 时拒绝 H_0 称为是**高度显著的**。

例2 某工厂生产的一种螺钉，标准要求长度是 32.5 毫米. 实际生产的产品，其长度 X 假定服从正态分布 $N(\mu,\sigma^2)$, 其中 σ^2 未知，现从该厂生产的一批产品中抽取 6 件, 得尺寸数据如下：

32.56, 29.66, 31.64, 30.00, 31.87, 31.03

问：这批产品是否合格?



解： 由已知,

设：这批产品(螺钉长度)的全体组成问题的总体为 X

则问题是要检验 $E(X)$ 是否为32.5.

…

第一步： 提出原假设和备择假设

$$H_0: \mu = 32.5 \quad \Leftrightarrow \quad H_1: \mu \neq 32.5$$

第二步： 因为已知 $X \sim N(\mu, \sigma^2)$, σ^2 未知.

故取检验统计量为：

$$t = \frac{\bar{X} - 32.5}{S/\sqrt{6}}$$

在 H_0 成立下求出它的分布为：

$$t = \frac{\bar{X} - 32.5}{S/\sqrt{6}} \sim t(5)$$

能衡量差异大小且分布已知

第三步：

对给定的显著性水平 $\alpha = 0.01$ 查 t 分布表得临界值：

$$t_{\alpha/2}(5) = t_{0.005}(5) = 4.0322$$

使得： $P\{|t| > t_{\alpha/2}(5)\} = \alpha$

即“ $|t| > t_{\alpha/2}(5)$ ”是一个小概率事件．

小概率事件在一次试验中基本上是不会发生．

故得否定域为：

$$C: \quad |t| > 4.0322$$

第四步：

将样本值代入，计算出统计量 t 的实测值：

$$|t| = 2.997 < 4.0322$$

没有落入拒绝域

故不能拒绝H_0，即应接受H_0

结论：可认为这批产品是合格的。

注：接受H_0这并不意味着H_0一定对，只是差异还不够显著，不足以否定H_0。

例3. 设某异常区磁场强度服从正态分布 $N(\mu_0,\sigma_0^2)$，由以前观察知道 $\mu_0=56$, $\sigma_0=20$，现有一台新型号的仪器，用它对该区进行磁测，抽取了41个点，其样本均值与方差为：

$$\bar{X}=61.1,\quad s=20$$

问：此仪器测出的结果是否符合要求?

解：以 μ,σ 分别表示用这台机器测出的异常区的磁场强度 X 的均值和均方差(标准差)。

根据长期实践的经验表明异常区磁场强度的标准差比较稳定，所以可设 $\sigma=20$,

于是：$X\sim N(\mu,20^2)$ 这里 μ 是未知的.

第一步： 提出假设： $H_0: \quad \mu = \mu_0 = 56$

$$H_1: \quad \mu \neq \mu_0 = 56$$

第二步： 由已知条件取检验统计量为：

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

第三步：

对给定的显著性水平 $\alpha = 0.05$ 查正态分布表得临界值： $k = z_{\frac{0.05}{2}} = z_{0.025} = 1.96$

使得：$$P_{\mu_0}\left\{\left|\frac{\bar{X}-\mu_0}{\sigma/\sqrt{n}}\right|\geq k\right\}=\alpha$$

即：$\left|\frac{\bar{X}-\mu_0}{\sigma/\sqrt{n}}\right|\geq k=1.96$ 是一个小概率事件.

故得否定域为： $C:\left|\frac{\bar{X}-\mu_0}{\sigma/\sqrt{n}}\right|\geq 1.96$

第四步：

将样本值代入，计算出统计量 U 的实测值：

$$u=\left|\frac{\bar{x}-\mu_0}{\sigma/\sqrt{n}}\right|=\left|\frac{61.1-56}{20/\sqrt{41}}\right|=\left|\frac{5.1}{3.125}\right|=1.632<1.96$$

没有落入拒绝域

故不能拒绝H_0，即应接受H_0

结论：

可认为这台仪器测出的结果是符合要求的。

即这台机器是基本正常的。

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

显著性水平 α

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

------犯第一类错误的概率，

C 为拒绝域。

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.）

作出决策

拒绝 H_0 还是接受 H_0

注: ▲ 备择假设 H_1 表示 μ 可能大于 μ_0 也可能小于 μ_0，故称其为 双边备择假设。从而对应的假设检验称为 双边假设检验。

▲ 拒绝域与临界点

(1)当统计量取某个区域 C 中的值时，拒绝原假设 H_0，则称区域 C 为 拒绝域。

(2) 拒绝域的边界点称为 临界点

▲ 单边检验

(1) 右边检验: H_0 : $\mu = \mu_0$ H_1 : $\mu > \mu_0$

(2) 左边检验: H_0 : $\mu = \mu_0$ H_1 : $\mu < \mu_0$

例如，在正态分布中针对显著性水平α，一般有：

当 $$\left|\frac{\overline{x}-\mu_0}{\sigma/\sqrt{n}}\right|<k$$ 则称 $\overline{x}$ 与 μ_0 的差异不显著

当 $$\left|\frac{\overline{x}-\mu_0}{\sigma/\sqrt{n}}\right|\geq k$$ 则称 $\overline{x}$ 与 μ_0 的差异显著

五. 假设检验问题的步骤

1. 根据实际问题要求，提出原假设H_0及备择假设H_1

2. 给定显著性水平α 及样本容量n

3. 确定检验统计量及拒绝域的形式

4. 按 $P(\text{拒绝}H_0 | H_0\text{为真}) = \alpha$ ，求出拒绝域

5. 取样本，根据样本观察值确定接受H_0还是拒绝H_0

例4　某编织物强力指标 X 的均值 $\mu_0 = 21$ 公斤。改进工艺后生产了一批编织物，今从中取 30 件，测得 $\bar{X} = 21.55$ 公斤。假设强力X 指标服从正态分布 $N(\mu, \sigma^2)$，且已知 $\sigma = 1.2$ 公斤。

问：在显著性水平 $\alpha = 0.01$ 下，新生产编织物比过去的编织物强力是否有提高？

解：提出假设：$H_0: \mu \le 21 \quad \Leftrightarrow \quad H_1: \mu > 21$

取统计量：$U = \dfrac{\bar{X} - 21}{\sigma/\sqrt{n}} \sim N(0,1)$

$\{U > u_{0.01}\}$ 是一小概率事件

否定域 C 为：$U > z_{0.01} = 2.33$

由已知，$\sigma = 1.2$，$n = 30$

并由样本值计算，得统计量 U 的实测值为：

$$u = 2.51 > 2.33$$

落入否定域

故拒绝原假设 H_0，可认为新生产编织物比过去的编织物强力是有提高的。

注：此时可能会犯第一类错误，但犯错误的概率不会超过 0.01.