第八章 假设检验

• 中山大学 • 计算机学院 • 郑培嘉 • Email：zhpj@mail.sysu.edu.cn

目录

1. 假设检验

假设检验

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 → 但有猜测怀疑，需要证实之时 → 用假设检验的方法来处理

假设检验

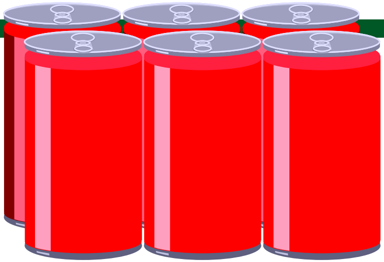
假设检验是统计推断的另一类重要组成部分. 它分为参数假设检验与非参数假设检验.

参数假设检验是对总体分布函数中的未知参数提出某种假设, 然后利用样本提供的信息对所提出的假设进行检验, 根据检验的结果对所提出的假设作出拒绝或接受的判断.

非参数假设检验是对总体分布函数的形式或总体的性质提出某种假设进行的检验.

参数假设检验与参数估计是从不同的角度推断总体分布中的某些参数, 参数检验解决定性问题, 参数估计解决定量问题.

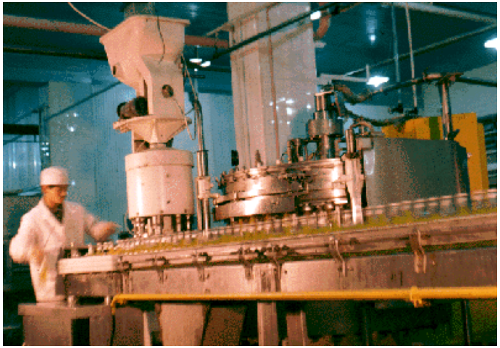
假设检验



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?

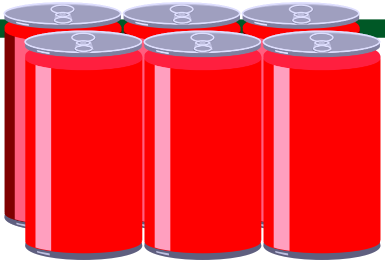


把每一罐都打开倒入量杯,
看看容量是否合于标准.

这样做显然不行!

假设检验

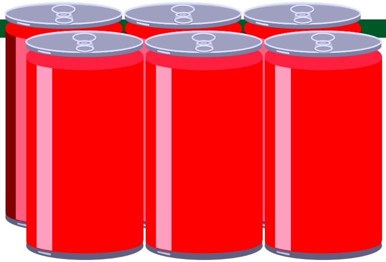
通常的办法是进行抽样检查.



如每隔1小时，抽查5罐，得5个容量的值$X1$， …，$X5$，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.

假设检验



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.

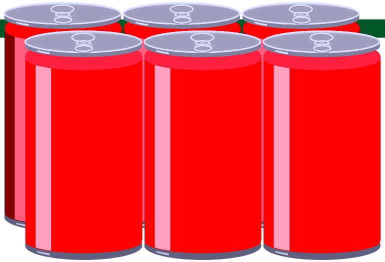
当然也不能总认为正常，有了问题不能及时发现，这也要造成损失.

如何处理这两者的关系，假设检验面对的就是这种矛盾.

假设检验

罐装可乐的容量按标准应在

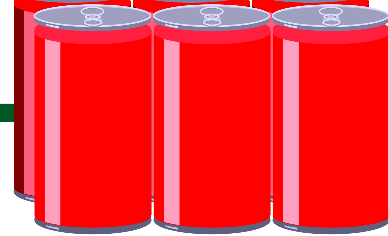
350毫升和360毫升之间.



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.

假设检验



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：

$$H_0: \mu = \mu_0 \ (\mu_0 = 355)$$

它的对立假设是：

$$H_1: \mu \neq \mu_0$$

在实际工作中，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，解消假设）；

称H_1为备选假设（或对立假设）.

假设检验

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，因此$\overline{X}$可以根据$\overline{X}$与μ_0的差距$|\overline{X}-\mu_0|$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，可以认为H_0不成立的；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？

假设检验

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”

假设检验

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．

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

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？

假设检验

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

在σ是一个常数情况下，选检验统计量$U = \frac{\bar{X}-\mu_0}{\sigma/\sqrt{n}} \sim \mathrm{N}(0,1)$

它能衡量差异大小$|\overline{X} - \mu_0|$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，使$P\{|U| > u_{\alpha/2}\} = \alpha$

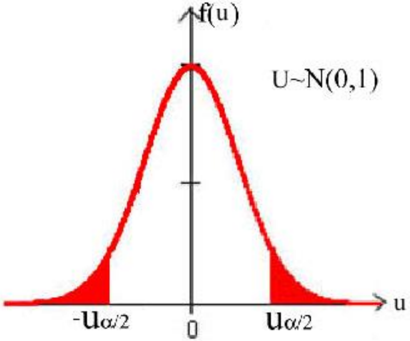
假设检验

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

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.

假设检验

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 .

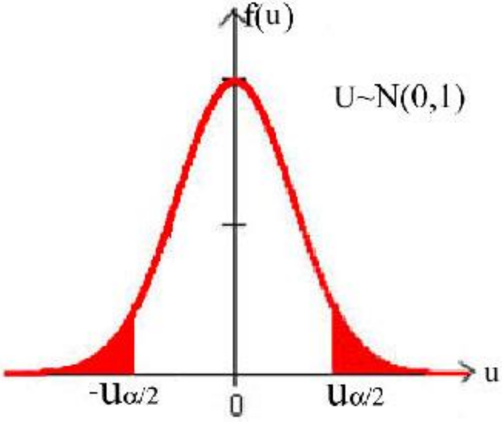
所以假设检验又叫"显著性检验"

假设检验

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.

假设检验

假设检验的一般步骤

在上面的例子的叙述中，我们已经初步介绍了假设检验的基本思想和方法．下面，我们再结合另一个例子，进一步说明假设检验的一般步骤．

例: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$N(\mu, \sigma^2)$, 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?

假设检验

分析：这批产品(螺钉长度)的全体组成问题的总体X. 现在要检验$E(X)$是否为32.5.

已知$X \sim N(\mu, \sigma^2)$, σ^2未知

第一步：提出原假设和备择假设

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

第二步：取一检验统计量，在H_0成立下求出它的分布

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

能衡量差异大小且分布已知

假设检验

第三步：对给定的显著性水平$\alpha = 0.01$，查表确定临界值

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

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

即“$|t| > t_{\alpha/2}(5)$ ”是一个小概率事件 .

得否定域　　W: $|t| > 4.0322$

第四步：将样本值代入算出统计量 t 的实测值,

$$|t| = 2.997 < 4.0322$$

没有落入拒绝域

故不能拒绝H_0 . 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 .

假设检验

假设检验的相关概念

（1） 原假设与备择假设

假设检验问题通常叙述为:在显著性水平α下, 检验假设

$$H_0: \ \mu = \mu_0, \quad H_1: \ \mu \neq \mu_0.$$

H_0称为原假设或零假设, H_1 称为备择假设.

（2） 否定域(拒绝域)

当检验统计量取某个区域W 中的值时，我们拒绝原假设H_0,

则称区域 W 为否定域(拒绝域).

假设检验

(3)．两类错误及记号

假设检验的依据是：小概率事件在一次试验中很难发生，但很难发生不等于不发生，因而假设检验所作出的结论有可能是错误的．这种错误有两类：

a) 当原假设H_0为真，观察值却落入否定域，而作出了拒绝H0的判断，称做第一类错误，又叫弃真错误，这类错误是“以真为假”. 犯第一类错误的概率是显著性水平α

b) 当原假设H_0不真，而观察值却落入接受域，而作出了接受H0 的判断，称做第二类错误，又叫取伪错误，这类错误是“以假为真”．

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

假设检验

犯第二类错误的概率记为

$$P\{当\ H_0\ 不真接受\ H_0\}\ 或\ P_{\mu\in H_1}\{接受\ H_0\}.$$

当样本容量n一定时，若减少犯第一类错误的概率，则犯第二类错误的概率往往增大.

若要使犯两类错误的概率都减小，除非增加样本容量.

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 .

假设检验

注: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.

(4). 显著性检验

只对犯第一类错误的概率加以控制, 而不考虑犯第二类错误的概率的检验, 称为显著性检验.