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

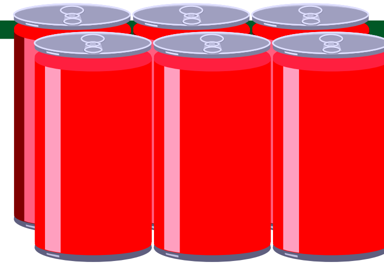
假设检验是统计推断的另一类重要组成部分. 它分为参数假设检验与非参数假设检验.

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.

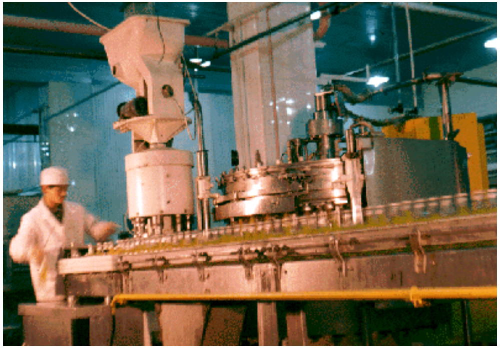
参数假设检验与参数估计是从不同的角度推断总体分布中的某些参数, 参数检验解决定性问题, 参数估计解决定量问题.

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.



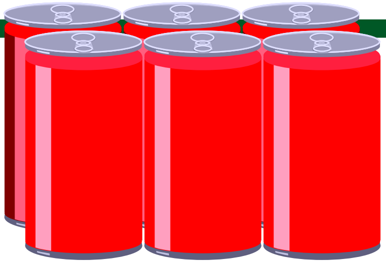
生产流水线上罐装可乐不断地封装，然后装箱外运. 怎么知道这批罐装可乐的容量是否合格呢？

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

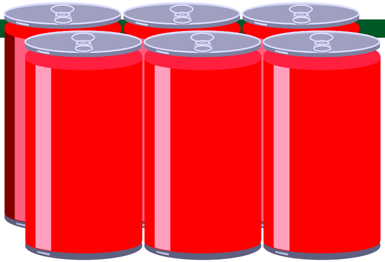
通常的办法是进行抽样检查.



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

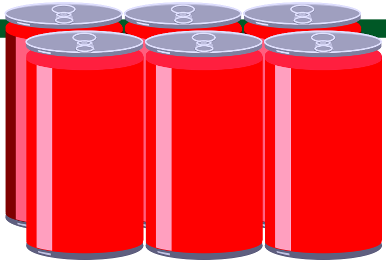


很明显，不能由5罐容量的数据，在把握不大的情况下就判断生产不正常，因为停产的损失是很大的.

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



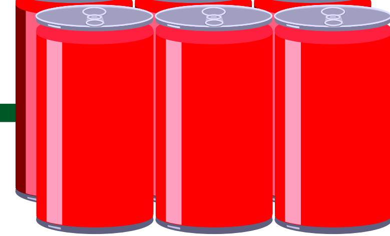
罐装可乐的容量按标准应在

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

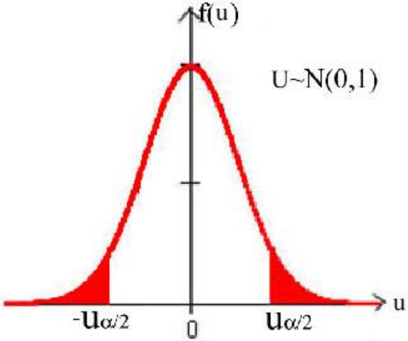
假设检验

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

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 .

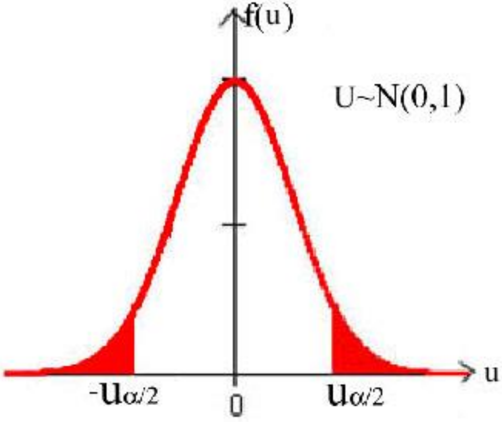
所以假设检验又叫"显著性检验"

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

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

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.

假设检验

（5）双边假设检验

在 H_0： $\mu = \mu_0$ 和 H_1： $\mu \neq \mu_0$ 中，备择假设 H_1 表示 μ 可能大于 μ_0，也可能小于 μ_0，称为双边备择假设，形如 H_0： $\mu = \mu_0$ ，H_1： $\mu \neq \mu_0$ 的假设检验称为双边假设检验.

（6）右边检验与左边检验

形如 H_0： $\mu \leq \mu_0$ ，H_1： $\mu > \mu_0$ 的假设检验称为右边检验.

形如 H_0： $\mu \geq \mu_0$ ，H_1： $\mu < \mu_0$ 的假设检验称为左边检验.

右边检验与左边检验统称为单边检验.

假设检验

（7） 单边检验的拒绝域

设总体$X \sim N\,(\mu,\ \sigma^2)$，$\sigma$为已知，$X_1, X_2, \cdots, X_n$是来自总体$X$的样本，给定显著性水平$\alpha$，

则 右边检验的拒绝域为

$$z = \frac{\bar{x} - \mu_0}{\sigma/\sqrt{n}} \geq z_\alpha,$$

左边检验的拒绝域为

$$z = \frac{\bar{x} - \mu_0}{\sigma/\sqrt{n}} \leq -z_\alpha.$$

假设检验

例：某工厂生产的固体燃料推进器的燃烧率服从正态分布 $N(\mu,\sigma^2),\mu=40cm/s,\sigma=2cm/s$.现在用新方法生产了一批推进器。从中随机取 $n=25$只，测得燃烧率的样本均值为$\bar{x}=41.25cm/s$.设在新方法下总体均方差仍为 2cm/s，问这批推进器的燃烧率是否较以往生产的推进器的燃烧率有显著的提高？取显著性水平$\alpha=0.05$.

假设检验

解:提出假设: H_0: $\mu < \mu_0 = 40, H_1 : \mu > \mu_0$

取统计量$U = \frac{\overline{x}-\mu_0}{\sigma/\sqrt{n}} \geq U_{0.05} = 1.645$

否定域为W :$U > U_{0.05}$=1. 645

代入$\sigma = 2, n = 25$，并由样本值计算得统计量U的实测值

$$U = 3.125 > 1.645$$

落入否定域

故拒绝H_0 ，即认为这批推进器的燃料率较以往生产的有显著的提高。

2. 正态总体均值的假设检验

正态总体均值的假设检验

抽样分布

χ^2分布：设$X_1,\cdots,X_n$相互独立, 且均服从正态分布$N(0,1)$,

则称随机变量$\chi^2=\sum_{i=1}^{n}X_i^2$服从自由度为$n$的$\chi^2$分布,

记为$\chi^2\sim\chi^2(n)$.

t分布：设$X\sim N(0,1)$，$Y\sim\chi^2(n)$,且X与Y相互独立，则称随

机变量$t=\frac{X}{\sqrt{Y/n}}$服从自由度为n的t分布，记为$t\sim t(n)$.

F分布：设$U\sim\chi^2(n_1)$，$V\sim\chi^2(n_2)$,U与V相互独立，则称随

机变量$F=\frac{U/n_1}{V/n_2}$服从自由度为（n_1,n_2）的分布, 记为

$F\sim F(n_1,n_2)$.

正态总体均值的假设检验

抽样分布相关定理

样本均值的分布

设$X \sim N(\mu, \sigma^2)$，$X_1, \cdots, X_n$是来自总体X的样本，则样本均值$\bar{X}$有$\bar{X} \sim N(\mu, \sigma^2/n)$.

样本方差、均值的分布

设$X_1, \cdots, x_n$是来自总体$N(\mu, \sigma^2)$的样本，$\bar{X}, S^2$分别是样本均值和样本方差，则有

(1) $\dfrac{(n-1)S^2}{\sigma^2} \sim \chi^2(n-1)$ (2)$\bar{X}$与S^2独立. (3)$\dfrac{\bar{X}-\mu}{S/\sqrt{n}} \sim t\,(n-1)$

正态总体均值的假设检验

两总体样本均值差、样本方差比的分布

设$X_1,\cdots,X_{n_1}$与$Y_1,\cdots,Y_{n_2}$分别来自总体$N(\mu_1,\sigma_1^2)$和$N(\mu_2,\sigma_2^2)$的样本，且这两个样本相互独立.$\bar{X},\bar{Y}$分别是这两个样本的样本均值；S_1^2,S_2^2分别是这两个样本的样本方差，则有

$$1、\frac{S_1^2/\sigma_1^2}{S_2^2/\sigma_2^2}\sim F(n_1-1,n_2-1)$$

$$2、\frac{\bar{X}-\bar{Y}-(\mu_1-\mu_2)}{\sqrt{\frac{(n_1-1)S_1^2+(n_2-1)S_2^2}{n_1+n_2-2}}\sqrt{\frac{1}{n_1}+\frac{1}{n_2}}}\sim t(n_1+n_2-2)$$

正态总体均值的假设检验

单个总体$N(\mu,\sigma^2)$均值μ的检验

1. 方差σ^2已知情况下

U检验法（Z检验法）

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

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

由$P\left\{\left|\dfrac{\bar{X}-\mu_0}{\frac{\sigma}{\sqrt{n}}}\right|\geq u_{\frac{\alpha}{2}}\right\}=\alpha$，确定拒绝域$|U|\geq u_{\frac{\alpha}{2}}$

如果统计量的观测值 $|U|=\left|\dfrac{\bar{x}-\mu_0}{\frac{\sigma}{\sqrt{n}}}\right|\geq u_{\frac{\alpha}{2}}$，则拒绝原假设；否则接受原假设

正态总体均值的假设检验

单个总体$N(\mu,\sigma^2)$均值μ的检验

2. 方差σ^2未知情况下

t检验法

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

构造T统计量$T = \frac{\bar{X}-\mu_0}{\frac{S}{\sqrt{n}}} \sim t\,(n-1)$

由$P\left\{\left|\frac{\bar{X}-\mu_0}{\frac{S}{\sqrt{n}}}\right| \geq t_{\frac{\alpha}{2}}(n-1)\right\} = \alpha$，确定拒绝域$|T| \geq t_{\frac{\alpha}{2}}(n-1)$

如果统计量的观测值$|t| = \left|\frac{\bar{x}-\mu_0}{\frac{S}{\sqrt{n}}}\right| \geq t_{\frac{\alpha}{2}}(n-1)$，则拒绝原假设

；否则接受原假设

正态总体均值的假设检验

例：某种元件的寿命X(以h计)服从正态分布$N(\mu,\sigma^2)$，μ,σ^2均未知，现测得16只元件的寿命，

159	280	101	212	224	379	179	264
222	362	168	250	149	260	485	170

问是否有理由认为元件的平均寿命大于225？

解：按题意需检验H_0: $\mu \le \mu_0 = 225, H_1$: $\mu > 225$

取$\alpha = 0.05$. 其拒绝域为$t = \frac{\bar{x}-\mu_0}{s/\sqrt{n}} \ge t_\alpha(n-1)$

$n = 16$, $t_\alpha(n-1) = 1.7531, \bar{x} = 241.5, s = 98.7259$代入

$$t = \frac{\bar{x} - \mu_0}{s/\sqrt{n}} = 0.6685 < 1.7531$$

t没有落在拒绝域中，故接受H_0，认为平均寿命不大于225h

正态总体均值的假设检验

两个总体$N(\mu,\sigma^2)$均值差的检验

设 μ_1,μ_2,σ^2 均未知(假设两总体方差相等)

假设 $H_0:\mu_1-\mu_2=\delta$, $H_1:\mu_1-\mu_2\neq\delta.$

构造T统计量$\dfrac{(\overline{X}-\overline{Y})-(\mu_1-\mu_2)}{\sqrt{\frac{1}{n}+\frac{1}{m}}S_w}\sim t(n+m-2)$, $S_w=\sqrt{\dfrac{(n-1)S_1^2+(m-1)S_2^2}{n+m-2}}$

由$P_{\mu_1-\mu_2=\delta}\left\{\left|\dfrac{(\overline{X}-\overline{Y})-\delta}{S_w\sqrt{\frac{1}{n_1}+\frac{1}{n_2}}}\right|\geq k\right\}=\alpha$得到$k=t_{\frac{\alpha}{2}}(n_1+n_2-2)$

确定拒绝域$|t|=\dfrac{\left|(\overline{x}-\overline{y})-\delta\right|}{s_w\sqrt{\frac{1}{n_1}+\frac{1}{n_2}}}\geq t_{\frac{\alpha}{2}}(n_1+n_2-2)$

正态总体均值的假设检验

例：用两种方法(A和B)测定冰自-0.72 度转变为0度的水融化热(单位：cal/g)，测得数据如下

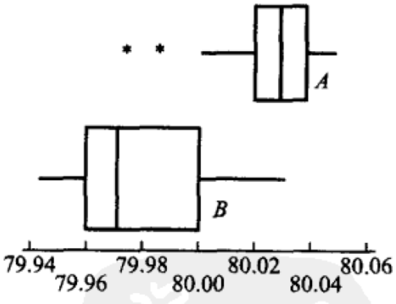
A：79.98 80.04 80.02 80.04 80.03 80.03 80.04 79.97 80.05 80.03 80.02 80.00 80.02

B：80.02 79.94 79.98 79.97 79.97 80.03 79.95 79.97

设两样本相互独立，且分别来自正态总体 $N(\mu_1,\sigma^2),N(\mu_2,\sigma^2),\mu_1,\mu_2,\sigma^2$均未知，检验假设(取显著性水平$\alpha=0.05$)

$$H_0: \mu_1-\mu_2\le 0, H_1: \mu_1-\mu_2>0$$

正态总体均值的假设检验



解：画出箱线图，可见是有明显差异的，

现在来检验上述假设

$$n_1 = 13, \overline{x_A} = 80.02, s_A^2 = 0.024^2$$

$$n_2 = 8, \overline{x_{\mathrm{B}}} = 79.98, s_{\mathrm{B}}^2 = 0.03^2$$

$$s_w^2 = \frac{12 \times s_A^2 + 7 \times s_{\mathrm{B}}^2}{19} = 0.0007178$$

$$|t| = \frac{\overline{x_A} - \overline{x_{\mathrm{B}}}}{s_w\sqrt{\frac{1}{13} + \frac{1}{8}}} = 3.33 > t_{0.05}(13 + 8 - 2) = 1.7291$$

故拒绝H_0，认为方法A比方法B测得的熔化热要大

正态总体均值的假设检验

基于成对数据的检验

有时为了比较两种产品、仪器、方法等的差异，我们常在相同的条件下作对比实验，得到一批成对的观察值，然后分析观察数据作出推断。这种方法常称为逐对比较法。

例：有两台光谱仪Ix，Iy 用于测量材料中某种金属的含量，为鉴定它们测量结果有无显著差异，制备9件试块（成分、金属含量、均匀性等各不相同）现分别用这两台仪器对每个试块测量一次，数据如下表

$x(\%)$	0.20	0.30	0.40	0.50	0.60	0.70	0.80	0.90	1.00
$y(\%)$	0.10	0.21	0.52	0.32	0.78	0.59	0.68	0.77	0.89
$d=x-y(\%)$	0.10	0.09	−0.12	0.18	−0.18	0.11	0.12	0.13	0.11

问：能否认为这两台仪器的测量结果有显著性差异($\alpha = 0.01$)

正态总体均值的假设检验

1. 表中每一行数据表示的是不同试块的测量在某一仪器下的测量结果，由于试块的成分、金属含量、均匀性等均不相同，因此不能看成来自于同一分布的样本数据；

2. 表中每一列数据表示的是同一个试块在两个不同仪器下的测量结果，因此这两数据间具有一定的相关性，不能认为是相互独立。

正态总体均值的假设检验

表中同一列中两数据的差异是由于两台仪器性能的差异引起的，因此我们以此分析仪器的差异；并且各列中的数据差异可以看成是相互独立的，且来自于同样的分布（有两机器差异决定）。

为此，设n对相互独立的观察结果$(X_1,Y_1)\ (X_1,Y_1) \dots (X_1,Y_1)$,

令 $D_1 = X_1 - Y_1;\ D_2 = X_2 - Y_2 \dots D_n = X_n - Y_n$,

则D_1，$D_2, \dots D_n$ 相互独立，且来自与同一分布。

正态总体均值的假设检验

假设$D_i \sim N(\mu_D, \sigma_D^2), i = 1,2,\ldots,n$，其中$\mu_D, \sigma_D^2$均未知，我们需要基于样本作出假设检验：

$$1. H_0: \mu_D = 0 \quad , \qquad H1: \mu_D \neq 0$$
$$2. H_0: \mu_D \leq 0 \quad , \qquad H1: \mu_D > 0$$
$$3. H_0: \mu_D \geq 0 \quad , \qquad H1: \mu_D < 0$$

记D_1，$D_2, \ldots D_n$ 的样本均值和样本方差的观察值为$\overline{d}, s_D^2$，按照单个正态总体的t检验，可得上述检验问题的拒绝域分别是

$$|t| = |\frac{\overline{d}}{s_D/\sqrt{n}}| \geq t_{\alpha/2}(n-1)$$
$$t = \frac{\overline{d}}{s_D/\sqrt{n}} \geq t_{\alpha}(n-1)$$
$$t = \frac{\overline{d}}{s_D/\sqrt{n}} \leq -t_{\alpha}(n-1)$$

正态总体均值的假设检验

例题中是第一种假设检验情况，$n = 9, t_{\frac{\alpha}{2}}(8) = t_{0.005}(8) = 3.3554$, 拒绝域为

$$|t| = \left|\frac{\overline{d}}{\frac{s_D}{\sqrt{n}}}\right| \geq 3.3554$$

带入观察值$\overline{d} = 0.06,\ s_D = 0.1227$

$$|t| = \left|\frac{0.06}{\frac{0.1227}{\sqrt{9}}}\right| = 1.467 < 3.3554$$

即|t|值不落在拒绝域上，因此接受H_0. 认为两台仪器的测量结果并没有显著差异。

3. 正态总体方差的假设检验

正态总体方差的假设检验

单个总体$N(\mu,\sigma^2)$方差σ^2的检验

1. 均值μ已知情况下

假设$H_0:\sigma^2=\sigma_0^2\,;H_1:\sigma^2\neq\sigma_0^2$

构造χ^2统计量$\chi^2=\frac{\sum_{i=1}^{n}(X_i-\mu)^2}{\sigma_0^2}\sim\chi^2(n)$

由$P\left\{\chi^2\leq\chi^2_{1-\frac{\alpha}{2}}(n)\right\}=\frac{\alpha}{2},P\left\{\chi^2\geq\chi^2_{\frac{\alpha}{2}}(n)\right\}=\frac{\alpha}{2}$，确定临界值

$\chi^2_{1-\frac{\alpha}{2}}(n),\chi^2_{\frac{\alpha}{2}}(n)$

如果统计量的观测值$\chi^2\geq\chi^2_{\frac{\alpha}{2}}(n)$或$\chi^2\leq\chi^2_{\frac{1-\alpha}{2}}(n)$，则拒绝原假设；否则接受原假设

正态总体方差的假设检验

单个总体$N(\mu,\sigma^2)$方差σ^2的检验

2. 均值μ未知情况下

χ^2检验法

由于S^2是σ^2的无偏估计，当H_0为真时，比值$\frac{S^2}{\sigma^2}$一般来说应在1附近摆动，而不应过分大于1或过分小于1。由于当 H0 为真时，

$$\frac{(n-1)S^2}{\sigma^2} \sim \chi^2(n-1)$$

我们取$\chi^2=\frac{(n-1)S^2}{{\sigma_0}^2}$作为检验统计量

正态总体方差的假设检验

由

$$P\left\{\chi^2 \le \chi^2_{1-\frac{\alpha}{2}}(n-1)\right\} = \frac{\alpha}{2}, P\left\{\chi^2 \ge \chi^2_{\frac{\alpha}{2}}(n-1)\right\} = \frac{\alpha}{2}$$

确定临界值 $\chi^2_{1-\frac{\alpha}{2}}(n-1), \chi^2_{\frac{\alpha}{2}}(n-1)$

拒绝域：

$$\chi^2 \ge \chi^2_{\frac{\alpha}{2}}(n-1) 或 \chi^2 \le \chi^2_{\frac{1-\alpha}{2}}(n-1)$$

则拒绝原假设；否则接受原假设

单边检验类似可得，可查表

正态总体均值、方差检验法（显著性水平α）

	原假设 H_0	检验统计量	备择假设 H_1	拒绝域
1	$\mu\leqslant\mu_0$ $\mu\geqslant\mu_0$ $\mu=\mu_0$ (σ^2 已知)	$Z=\dfrac{\overline{X}-\mu_0}{\sigma/\sqrt{n}}$	$\mu>\mu_0$ $\mu<\mu_0$ $\mu\neq\mu_0$	$z\geqslant z_\alpha$ $z\leqslant -z_\alpha$ $\lvert z\rvert\geqslant z_{\alpha/2}$
2	$\mu\leqslant\mu_0$ $\mu\geqslant\mu_0$ $\mu=\mu_0$ (σ^2 未知)	$t=\dfrac{\overline{X}-\mu_0}{S/\sqrt{n}}$	$\mu>\mu_0$ $\mu<\mu_0$ $\mu\neq\mu_0$	$t\geqslant t_\alpha(n-1)$ $t\leqslant -t_\alpha(n-1)$ $\lvert t\rvert\geqslant t_{\alpha/2}(n-1)$
3	$\mu_1-\mu_2\leqslant\delta$ $\mu_1-\mu_2\geqslant\delta$ $\mu_1-\mu_2=\delta$ (σ_1^2,σ_2^2 已知)	$Z=\dfrac{\overline{X}-\overline{Y}-\delta}{\sqrt{\dfrac{\sigma_1^2}{n_1}+\dfrac{\sigma_2^2}{n_2}}}$	$\mu_1-\mu_2>\delta$ $\mu_1-\mu_2<\delta$ $\mu_1-\mu_2\neq\delta$	$z\geqslant z_\alpha$ $z\leqslant -z_\alpha$ $\lvert z\rvert\geqslant z_{\alpha/2}$

正态总体均值、方差检验法(显著性水平α)

	原假设 H_0	检验统计量	备择假设 H_1	拒绝域
4	$\mu_1-\mu_2\leqslant\delta$ $\mu_1-\mu_2\geqslant\delta$ $\mu_1-\mu_2=\delta$ ($\sigma_1^2=\sigma_2^2=\sigma^2$ 未知)	$t=\dfrac{\overline{X}-\overline{Y}-\delta}{S_w\sqrt{\dfrac{1}{n_1}+\dfrac{1}{n_2}}}$ $S_w^2=\dfrac{(n_1-1)S_1^2+(n_2-1)S_2^2}{n_1+n_2-2}$	$\mu_1-\mu_2>\delta$ $\mu_1-\mu_2<\delta$ $\mu_1-\mu_2\neq\delta$	$t\geqslant t_\alpha(n_1+n_2-2)$ $t\leqslant -t_\alpha(n_1+n_2-2)$ $\|t\|\geqslant t_{\alpha/2}(n_1+n_2-2)$
5	$\sigma^2\leqslant\sigma_0^2$ $\sigma^2\geqslant\sigma_0^2$ $\sigma^2=\sigma_0^2$ (μ 未知)	$\chi^2=\dfrac{(n-1)S^2}{\sigma_0^2}$	$\sigma^2>\sigma_0^2$ $\sigma^2<\sigma_0^2$ $\sigma^2\neq\sigma_0^2$	$\chi^2\geqslant\chi_\alpha^2(n-1)$ $\chi^2\leqslant\chi_{1-\alpha}^2(n-1)$ $\chi^2\geqslant\chi_{\alpha/2}^2(n-1)$或 $\chi^2\leqslant\chi_{1-\alpha/2}^2(n-1)$
6	$\sigma_1^2\leqslant\sigma_2^2$ $\sigma_1^2\geqslant\sigma_2^2$ $\sigma_1^2=\sigma_2^2$ (μ_1,μ_2 未知)	$F=\dfrac{S_1^2}{S_2^2}$	$\sigma_1^2>\sigma_2^2$ $\sigma_1^2<\sigma_2^2$ $\sigma_1^2\neq\sigma_2^2$	$F\geqslant F_\alpha(n_1-1,n_2-1)$ $F\leqslant F_{1-\alpha}(n_1-1,n_2-1)$ $F\geqslant F_{\alpha/2}(n_1-1,n_2-1)$或 $F\leqslant F_{1-\alpha/2}(n_1-1,n_2-1)$
7	$\mu_D\leqslant 0$ $\mu_D\geqslant 0$ $\mu_D=0$ (成对数据)	$t=\dfrac{\overline{D}-0}{S_D/\sqrt{n}}$	$\mu_D>0$ $\mu_D<0$ $\mu_D\neq 0$	$t\geqslant t_\alpha(n-1)$ $t\leqslant -t_\alpha(n-1)$ $\|t\|\geqslant t_{\alpha/2}(n-1)$

正态总体方差的假设检验

两个总体$N(\mu,\sigma^2)$方差σ^2的检验

$\mu_1,\mu_2,\sigma_1^2,\sigma_2^2$均未知情况下

F检验法

检验假设

$$H_0:\sigma_1^2\le\sigma_2^2\,;H_1:\sigma_1^2>\sigma_2^2$$

当H_1为真时，观察值$\frac{S_1^2}{S_2^2}$有偏大的趋势，故拒绝域形式$\frac{S_1^2}{S_2^2}\ge \mathrm{k}$

构造统计量$F=\frac{S_1^2}{S_2^2}$，有$\frac{S_1^2/\sigma_1^2}{S_2^2/\sigma_2^2}\sim F\,(n-1,m-1)$

由$P_{\sigma_1^2\le\sigma_2^2}^2\left\{\frac{S_1^2}{S_2^2}\ge k\right\}\le P_{\sigma_1^2\le\sigma_2^2}^2\left\{\frac{\frac{S_1^2}{S_2^2}}{\frac{\sigma_1^2}{\sigma_2^2}}\ge k\right\}=\alpha$(因为$\sigma_1^2/\sigma_2^2\le 1$)，

$k=F_\alpha\,(n-1,m-1)$，拒绝域为$F=\frac{s_1^2}{s_2^2}\ge F_a(\mathrm{n}-1,\mathrm{m}-1)$.

6. 分布拟合检验

分布拟合检验

在前面的课程中，我们已经了解了假设检验的基本思想，并讨论了当总体分布为正态时，关于其中未知参数的假设检验问题．然而可能遇到这样的情形，总体服从何种理论分布并不知道，要求我们直接对总体分布提出一个假设．

解决这类问题的工具是英国统计学家K.皮尔逊在1900年发表的一篇文章中引进的所谓χ^2检验法.

这是一项很重要的工作，不少人把它视为近代统计学的开端.

χ^2检验法是在总体X的分布未知时，根据来自总体的样本，检验关于总体分布的假设的一种检验方法.

分布拟合检验

使用χ^2检验对总体分布进行检验时，我们先提出原假设:

$$H_0\text{：总体}X\text{的分布函数为}F(x)$$

然后根据样本的经验分布和所假设的理论分布之间的吻合程度来决定是否接受原假设.

这种检验通常称作拟合优度检验，它是一种非参数检验.

分布拟合检验

分布拟合的χ^2检验的基本原理和步骤如下：

1. 将总体X的取值范围分成k个互不重迭的小区间，记作 $A_1, A_2, \dots, A_k$.

2. 把落入第i个小区间A_i的样本值的个数记作f_i，称为实测频数. 所有实测频数之和$f_1 + f_2 + \cdots + f_k$等于样本容量n.

3. 根据所假设的理论分布，可以算出总体X的值落入每个A_i的概率p_i，于是np_i就是落入A_i的样本值的理论频数.

分布拟合检验

实测频数

$$f_i - np_i$$

理论频数

标志着经验分布与理论分布之间的差异的大小.

皮尔逊引进如下统计量表示经验分布

与理论分布之间的差异:

$$\chi^2 = \sum_{i=1}^{k} \frac{(f_i - np_i)^2}{np_i}$$

在理论分布已知的条件下, np_i是常量

分布拟合检验

皮尔逊证明了如下定理:

若原假设中的理论分布$F(x)$已经完全给定，那么当$n\to\infty$时，统计量

$$\chi^2=\sum_{i=1}^{k}\frac{(f_i-np_i)^2}{np_i}=\sum_{i=1}^{k}\frac{f_i^2}{np_i}-n$$

的分布渐近k-1个自由度的χ^2分布.

如果理论分布$F(x)$中有r个未知参数需用相应的估计量来代替，那么当$n\to\infty$时，统计量χ^2的分布渐近服从$(k-r-1)$个自由度的χ^2分布.

分布拟合检验

为了便于理解，我们对定理作一点直观的说明.

在理论分布$F(x)$完全给定的情况下，每个p_i 都是确定的常数.

由棣莫佛一拉普拉斯中心极限定理，当n充分大时，实测频数f_i渐近正态，

因此

$$\chi^2 = \sum_{i=1}^{k} \frac{(f_i - np_i)^2}{np_i}$$

是k个近似正态的变量的平方和.

这些变量之间存在着一个制约关系：$\sum_{i=1}^{k} \frac{\sqrt{p_i}(n_i - np_i)}{\sqrt{np_i}} = 0$

故统计量χ^2渐近服从(k−1)个自由度的χ^2分布。

分布拟合检验

在$F(x)$尚未完全给定的情况下，每个未知参数用相应的估计量代替，就相当于增加一个制约条件，因此，自由度也随之减少一个.

若有r个未知参数需用相应的估计量来代替，自由度就减少r个.

故统计量χ^2渐近服从$(k-r-1)$个自由度的χ^2分布。

分布拟合检验

根据这个定理，对给定的显著性水平α，查χ^2分布表可得临界值χ_α^2，使得

$$P(\chi^2 > \chi_\alpha^2) = \alpha$$

得拒绝域:

$$\chi^2 > \chi_\alpha^2(k-1)$$（不需估计参数）

$$\chi^2 > \chi_\alpha^2(k-r-1)$$（估计r个参数）

如果根据所给的样本值$X_1, X_2, \dots, X_n$算得统计量χ^2的实测值落入拒绝域，则拒绝原假设，否则就认为差异不显著而接受原假设.

分布拟合检验

注：皮尔逊定理是在n无限增大时推导出来的，因而在使用时要注意n要足够大，以及np_i 不太小这两个条件.

根据计算实践，要求n不小于50，以及np_i都不小于5．否则应适当合并区间，使np_i满足这个要求 .

分布拟合检验

例：下表列出了某一地区在下接的一个月中由100个气象站报告的雷暴雨的次数

i	0	1	2	3	4	5	$\geqslant 6$
f_i	22	37	20	13	6	2	0
A_i	A_0	A_1	A_2	A_3	A_4	A_5	A_6

其中f_i是报告雷暴雨次数为i的气象站数，使用χ^2拟合检验法检验雷暴雨次数X是否服从均值为$\lambda = 1$的泊松分布$\alpha = 0.05$

解：按题意需检验假设

$$H_0: P\{X = i\} = \frac{\lambda^i \mathrm{e}^{-\lambda}}{i!} = \frac{\mathrm{e}^{-1}}{i!}, i = 0,1,\cdots$$

分布拟合检验

$$H_0: P\{X=i\}=\frac{\lambda^i e^{-\lambda}}{i!}=\frac{e^{-1}}{i!}, i=0,1,\cdots$$

在H_0下X所有可能取的值为$\Omega=\{0,1,2,\ldots\}$, 将Ω分成如表所示的两两不相交的子集$A_0, A_1 \ldots A_6$, 则有

$$p_i=P\{X=i\}=\frac{e^{-1}}{i!}, i=0,1,\cdots,5.$$

例如

$$p_0=P\{X=0\}=e^{-1}=0.36788,$$

$$p_3=P\{X=3\}=\frac{e^{-1}}{3!}=0.06131,$$

$$p_6=P\{X\geqslant 6\}=1-\sum_{i=0}^{5}p_i=0.059.$$

$n=100.$

A_i	f_i	p_i	np_i		$f_i^2/(np_i)$
$A_0:\{X=0\}$	22	e^{-1}	36.788		13.16
$A_1:\{X=1\}$	37	e^{-1}	36.788		37.21
$A_2:\{X=2\}$	20	$e^{-1}/2$	18.394		21.75
$A_3:\{X=3\}$	13	$e^{-1}/6$	6.131	8.03	54.92
$A_4:\{X=4\}$	6	$e^{-1}/24$	1.533		
$A_5:\{X=5\}$	2	$e^{-1}/120$	0.307		
$A_6:\{X\geqslant 6\}$	0	$1-\sum_{i=0}^{5}p_i$	0.059		

$\sum=127.04$

YSU

分布拟合检验

对有些$np_i < 5$的组适当合并

A_i	f_i	p_i	np_i		$f_i^2/(np_i)$
$A_0:\{X=0\}$	22	e^{-1}	36.788		13.16
$A_1:\{X=1\}$	37	e^{-1}	36.788		37.21
$A_2:\{X=2\}$	20	$e^{-1}/2$	18.394		21.75
$A_3:\{X=3\}$	13	$e^{-1}/6$	6.131	8.03	54.92
$A_4:\{X=4\}$	6	$e^{-1}/24$	1.533		
$A_5:\{X=5\}$	2	$e^{-1}/120$	0.307		
$A_6:\{X\geqslant 6\}$	0	$1-\sum_{i=0}^{5}p_i$	0.059		

$\sum=127.04$

合并组后，$k = 4$，自由度为3，将数据代入公式得到

$$\chi^2 = 27.04 > 7.815 = \chi^2_{0.05}(3)$$

故落在拒绝域内，在显著性水平0.05下拒绝H_0，认为样本不是来自均值$\lambda = 1$的泊松分布

8. 假设检验问题的p值法

假设检验问题的p值法

假设检验的结论通常是简单的：在给定的显著水平下，不是拒绝原假设就是保留原假设。然而有时也会出现这样的情况：在一个较大的显著水平$\alpha = 0.05$下得到拒绝原假设的结论，而在一个较小的显著水平$\alpha = 0.01$下却会得到相反的结论。

这种情况在理论上很容易解释：

假设检验问题的p值法

因为显著水平变小后会导致检验的拒绝域变小，于是原来落在拒绝域中的观测值就可能落入接受域。

但这种情况在应用中会带来一些麻烦：假如这时一个人主张选择显著水平$\alpha = 0.05$，而另一个人主张选$\alpha = 0.01$，则第一个人的结论是拒绝H_0，而后一个人的结论是接受H_0，我们该如何处理这一问题呢？

假设检验问题的p值法

定义:在一个假设检验问题中，利用观测值能够做出拒绝原假设的**最小显著性水平称为检验的p值。**

引进检验的p值的概念有明显的好处:

第一，它比较客观，避免了事先确定显著水平;

其次，由检验的p 值与人们心目中的显著性水平进行比较可以很容易作出检验的结论:

- 如果$\alpha \geq p$，则在显著性水平α下拒绝H_0;
- 如果$\alpha < p$，则在显著性水平α下保留H_0.

p值在应用中很方便，如今的统计软件中对检验问题一般都会给出检验的p 值。

假设检验问题的p值法

例：用p值法检验本章2例2的检验问题

例 2 公司从生产商购买牛奶。公司怀疑生产商在牛奶中掺水以谋利。通过测定牛奶的冰点，可以检验出牛奶是否掺水。天然牛奶的冰点温度近似服从正态分布，均值 $\mu_0=-0.545$ ℃，标准差 $\sigma=0.008$ ℃。牛奶掺水可使冰点温度升高而接近于水的冰点温度(0 ℃)。测得生产商提交的 5 批牛奶的冰点温度，其均值为 $\bar{x}=-0.535$ ℃，问是否可以认为生产商在牛奶中掺了水？取 $\alpha=0.05$.

$$H_0: \mu \le \mu_0 = -0.545, H_1: \mu > \mu_0$$

解：用Z检验法，检验统计量$Z=\frac{\bar{X}-\mu}{\sigma/\sqrt{n}}$的观察值为

$$z_0=\frac{-0.535-(-0.545)}{0.008/\sqrt{5}}=2.7955$$

p值=$P\{Z>=2.7955\}=1-\phi(2.7955)=0.0026$

$\alpha>p$, 故拒绝H_0.

THANKS

谢谢！