第一节 假设检验

一、假设检验的基本原理

二、假设检验的相关概念

三、假设检验的一般步骤

四、小结

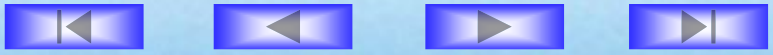
一、假设检验的基本原理

在总体的分布函数完全未知或只知其形式、但不知其参数的情况下，为了推断总体的某些性质，提出某些关于总体的假设.

例如，提出总体服从泊松分布的假设；又如，对于正态总体提出数学期望等于 μ_0 的假设等.

假设检验就是根据样本对所提出的假设作出判断：是接受，还是拒绝.

假设检验问题是作出这一决策的过程.



假设检验问题是统计推断的另一类重要问题.

如何利用样本值对一个具体的假设进行检验?

通常借助于直观分析和理论分析相结合的做法, 其基本原理就是人们在实际问题中经常采用的所谓实际推断原理:“**一个小概率事件在一次试验中几乎是不可能发生的**”.

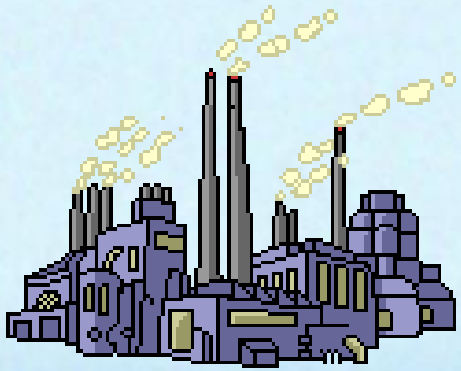
下面结合实例来说明假设检验的基本思想.

例1　某车间用一台包装机包装葡萄糖, 袋装糖的净重是一个随机变量，它服从正态分布. 当机器正常时, 其均值为0.5kg, 标准差为0.015kg. 某日开工后为检验包装机是否正常，随机地抽取它所包装的糖9袋，称得净重为(kg):

0.497　0.506　0.518　0.524　0.498　0.511

0.520　0.515　0.512,

问机器是否正常?



分析 以μ和σ分别表示这一天袋装糖的净重总体X的均值和标准差,

由长期实践表明标准差比较稳定, 我们就设$\sigma = 0.015$,于是$X \sim N(\mu, 0.015^2)$,这里μ未知.

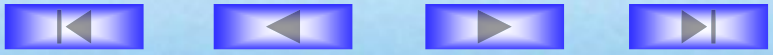
问题 问题是根据样本值判断$\mu = 0.5$还是$\mu \neq 0.5$.**为此,我们提出两个相互对立的假设**

$$H_0 : \mu = \mu_0 = 0.5$$

和

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

然后,我们给出一个合理的法则,根据这一法

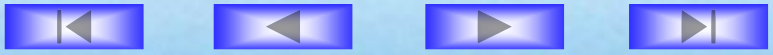


则,利用已知样本作出决策是接受假设H_0(即拒绝假设H_1), 还是拒绝假设H_0(即接受假设H_1). 如果作出的决策是接受H_0,则认为$\mu = \mu_0$, **即认为机器工作是正常的, 否则, 认为是不正常的.**

由于要检验的假设涉及总体均值μ,可借助样本均值$\overline{X}$进行判断.

$\overline{X}$ 是 μ的无偏估计, H_0为真, $|\overline{x} - \mu_0|$ 不应太大, 当H_0为真时

$$\frac{\overline{X} - \mu_0}{\sigma / \sqrt{n}} \sim N(0,1).$$



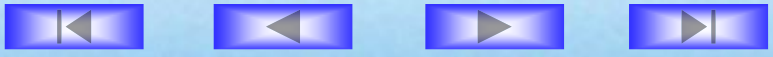
衡量 $|\bar{x}-\mu_0|$ 的大小可归结为衡量 $\dfrac{|\bar{x}-\mu_0|}{\sigma/\sqrt{n}}$ 的大小. 适当选定一正数 k,

当观测值 $\bar{x}$ 满足 $\dfrac{|\bar{x}-\mu_0|}{\sigma/\sqrt{n}} \geq k$ 时，拒绝假设 H_0,

反之， 若 $\dfrac{|\bar{x}-\mu_0|}{\sigma/\sqrt{n}} < k$ 时, 就接受假设 H_0.

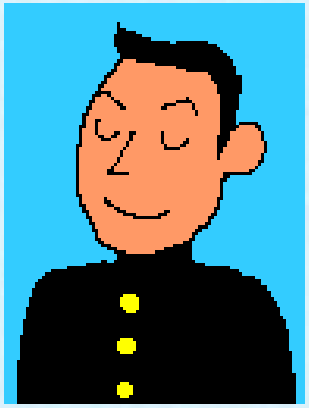
由于当 H_0 为真时， $Z=\dfrac{\bar{X}-\mu_0}{\sigma/\sqrt{n}} \sim N(0,1)$,

由标准正态分布分位点的定义得： $k=z_{\alpha/2}$,



当$\dfrac{|\bar{x}-\mu_0|}{\sigma/\sqrt{n}} \geq z_{\alpha/2}$时，拒绝$H_0$，$\dfrac{|\bar{x}-\mu_0|}{\sigma/\sqrt{n}} < z_{\alpha/2}$时，接受$H_0$.

假设检验过程如下:



在实例中若取定$\alpha = 0.05$,

则$k = z_{\alpha/2} = z_{0.025} = 1.96$,

又已知$n = 9$, $\sigma = 0.015$,

由样本算得 $\bar{x} = 0.511$, 即有 $\dfrac{|\bar{x}-\mu_0|}{\sigma/\sqrt{n}} = 2.2 > 1.96$,

于是拒绝假设H_0,认为包装机工作不正常.

以上所采取的检验法是符合实际推断原理的.

因通常α总是取得很小, 一般取 $\alpha = 0.01, 0.05$.

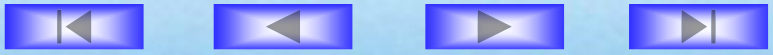
因而当H_0为真, 即$\mu = \mu_0$时, $\left\{ \dfrac{\overline{X} - \mu_0}{\sigma / \sqrt{n}} \geq z_{\alpha/2} \right\}$

是一个小概率事件, 根据实际推断原理, 就可以认

为如果H_0为真, 由一次试验得到满足不等式

$\left| \dfrac{\overline{x} - \mu_0}{\sigma / \sqrt{n}} \right| \geq z_{\alpha/2}$的观察值$\overline{x}$, 几乎是不会发生的.

在一次观测中竟出 现了满足 $\left|\dfrac{\bar{x}-\mu_0}{\sigma/\sqrt{n}}\right| \geq z_{\alpha/2}$的$\bar{x}$，我们有理由怀疑原来的假设$H_0$的正确性，因而拒绝$H_0$. 若出现观测值 $\bar{x}$ 满足不等式 $\left|\dfrac{\bar{x}-\mu_0}{\sigma/\sqrt{n}}\right| < z_{\alpha/2}$，则没有理由拒绝假设$H_0$，因而只能接受$H_0$.



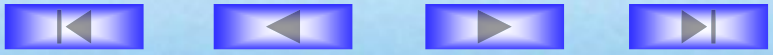
二、假设检验的相关概念

1. 显著性水平

如果$|z|=\left|\dfrac{\bar{x}-\mu_0}{\sigma/\sqrt{n}}\right|\geq k$, 则称$\bar{x}$与$\mu_0$的差异是显著的,则我们拒绝 H_0; 反之,如果$|z|=\left|\dfrac{\bar{x}-\mu_0}{\sigma/\sqrt{n}}\right|<k$,则称$\bar{x}$与$\mu_0$的差异是不显著的, 则我们接受 H_0.

数 α 称为显著性水平.

上述关于$\bar{x}$与 μ_0 有无显著差异的判断是在显著性水平α下做出的,则我们接受H_0.



2. 检验统计量

统计量 $Z=\dfrac{\overline{X}-\mu_0}{\sigma/\sqrt{n}}$ 称为检验统计量.

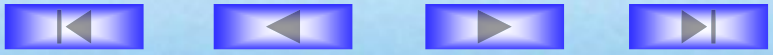
3. 原假设与备择假设

前面的检验问题通常叙述成: 在显著性水平α下, 检验假设

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

也常说成"在显著性水平α下, 针对 H_1检验 H_0".

H_0称为原假设或零假设, H_1 称为备择假设.



4. 拒绝域与临界点

当检验统计量取某个区域C中的值时, 我们拒绝原假设H_0,则称区域C为拒绝域, 拒绝域的边界点称为临界点.

如在上例中,

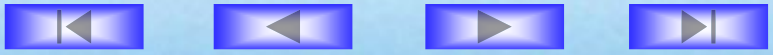
拒绝域为 $|z| \geq z_{\alpha/2}$,

临界点为 $z = z_{\alpha/2}$, $z = -z_{\alpha/2}$.

5. 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:

(1) 当原假设H_0为真，观察值却落入拒绝域，而作出了拒绝H_0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.



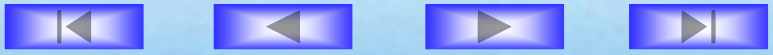
(2) 当原假设 H_0 不真，而观察值却落入接受域，而作出了接受 H_0 的判断，称做**第二类错误**，又叫**取伪错误**，这类错误是“以假为真”.

犯第II类错误的概率记为

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

一般来说，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.



6. 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.

7. 双边备择假设与双边假设检验

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.

8. 右边检验与左边检验

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

形如 $H_0:\mu=\mu_0$，$H_1:\mu\leq\mu_0$ 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.

9. 单边检验的拒绝域

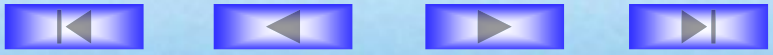
设总体$X \sim N(\mu, \sigma^2)$, σ为已知, $X_1, X_2, \cdots, X_n$是来自总体X的样本, 给定显著性水平α.

则: 右边检验的拒绝域为$z = \dfrac{\bar{x} - \mu_0}{\sigma / \sqrt{n}} \geq z_\alpha$,

左边检验的拒绝域为$z = \dfrac{\bar{x} - \mu_0}{\sigma / \sqrt{n}} \leq -z_\alpha$.

证明 (1)右边检验

因H_0中的全部μ都比H_1中的μ要小,当H_1为真时, 观察值$\bar{x}$往往偏大,



因此拒绝域的形式为 $\bar{x} \geq k$，(k是某一正常数)

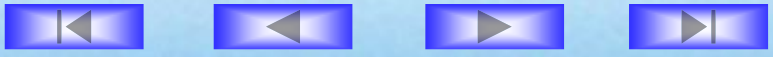
由 $P\{H_0 \text{为真拒绝} H_0\} = P_{\mu \in H_0}\{\bar{X} \geq k\}$

$$= P_{\mu \leq \mu_0}\left\{\frac{\bar{X} - \mu_0}{\sigma/\sqrt{n}} \geq \frac{k - \mu_0}{\sigma/\sqrt{n}}\right\}$$

$$\leq P_{\mu \leq \mu_0}\left\{\frac{\bar{X} - \mu}{\sigma/\sqrt{n}} \geq \frac{k - \mu_0}{\sigma/\sqrt{n}}\right\}$$

上式不等号成立的原因:

因为 $\mu \leq \mu_0$，所以 $\dfrac{\bar{X} - \mu}{\sigma/\sqrt{n}} \geq \dfrac{\bar{X} - \mu_0}{\sigma/\sqrt{n}}$，



事件 $\left\{\dfrac{\overline{X}-\mu_0}{\sigma/\sqrt{n}} \geq \dfrac{k-\mu_0}{\sigma/\sqrt{n}}\right\} \subset \left\{\dfrac{\overline{X}-\mu}{\sigma/\sqrt{n}} \geq \dfrac{k-\mu_0}{\sigma/\sqrt{n}}\right\}.$

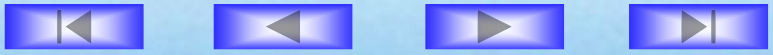
要控制 $P\{H_0 \text{ 为真拒绝 } H_0\} \leq \alpha$，只需令

$$P_{\mu \leq \mu_0}\left\{\frac{\overline{X}-\mu}{\sigma/\sqrt{n}} \geq \frac{k-\mu_0}{\sigma/\sqrt{n}}\right\} = \alpha.$$

由于 $$\frac{\overline{X}-\mu}{\sigma/\sqrt{n}} \sim N(0,1),$$

所以 $k = \mu_0 + \dfrac{\sigma}{\sqrt{n}} z_\alpha,$ $\dfrac{k-\mu_0}{\sigma/\sqrt{n}} = z_\alpha,$

故右边检验的拒绝域为



$$\bar{x} \geq \mu_0 + \frac{\sigma}{\sqrt{n}} z_\alpha,$$

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

类似证，

左边检验的拒绝域为 $z = \dfrac{\bar{x} - \mu_0}{\sigma / \sqrt{n}} \leq -z_\alpha.$

三、假设检验的一般步骤

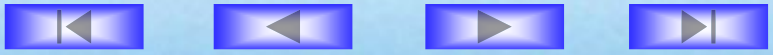
1. 根据实际问题的要求,提出原假设 H_0 及备择假设 H_1;

2. 给定显著性水平 α 以及样本容量 n;

3. 确定检验统计量以及拒绝域形式;

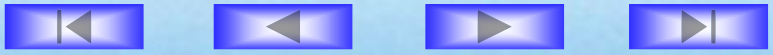
4. 按 $P\{H_0$ 为真拒绝 $H_0\} \leq \alpha$ 求出拒绝域;

5. 取样,根据样本观测值确定接受还是拒绝 H_0.



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$\mu_0 = -0.545°C$, 标准差$\sigma = 0.008°C$.

牛奶掺水可使冰点温度升高而接近于水的冰点温度$(0°C)$. 测得生产商提交的5批牛奶的冰点温度, 其均值$\bar{x} = -0.535°C$, 问是否可以认为生产商在牛奶中掺了水? 取$\alpha = 0.05$.



解　按题意需检验假设

H_0：　$\mu \le \mu_0 = -0.545$　(即设牛奶未掺水)

H_1：　$\mu > \mu_0$　(即设牛奶已掺水)

这是右边检验问题，其拒绝域为

$$z = \frac{\bar{x} - \mu_0}{\sigma/\sqrt{n}} \ge z_{0.05} = 1.645.$$

现在　$z = \dfrac{-0.535 - (-0.545)}{0.008/\sqrt{5}} = 2.7951 > 1.645,$

z的值落在拒绝域中，所以我们在显著性水平

$\alpha = 0.05$下拒绝H_0，即认为牛奶商在牛奶中掺了水.

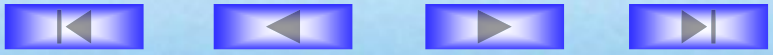
补充例题

四、小结

假设检验的基本原理、相关概念和一般步骤.

假设检验的两类错误

真实情况(未知)	所　作　决　策	
	接受 H_0	拒绝 H_0
H_0 为真	正确	犯第I类错误
H_0 不真	犯第II类错误	正确



第二节 正态总体均值的假设检验

一、单个总体均值 μ 的检验

二、两个总体均值差的检验(t 检验)

三、基于成对数据的检验(t 检验)

四、小结

一、单个总体$N(\mu,\sigma^2)$均值μ 的检验

1. σ^2 为已知, 关于μ的检验(Z 检验)

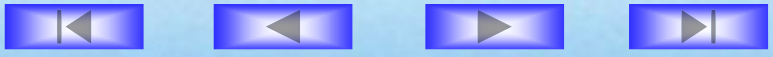
在上节中讨论过正态总体 $N(\mu,\sigma^2)$,

当σ^2为已知时, 关于$\mu=\mu_0$的检验问题 :

(1) 假设检验 $H_0:\mu=\mu_0$, $H_1:\mu\neq\mu_0$;

(2) 假设检验 $H_0:\mu\leq\mu_0$, $H_1:\mu>\mu_0$;

(3) 假设检验 $H_0:\mu\geq\mu_0$, $H_1:\mu<\mu_0$.



在这些检验问题中，我们都是利用统计量

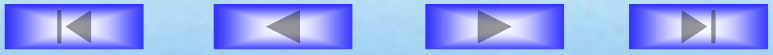
$Z=\dfrac{\overline{X}-\mu_0}{\sigma/\sqrt{n}}$ 来确定拒绝域的，这种检验法称为 Z 检验法.

一个有用的结论

当显著性水平均为 α 时，

检验问题 $H_0:\mu\le\mu_0$，　$H_1:\mu>\mu_0$ 和检验问题

$H_0:\mu=\mu_0$，　$H_1:\mu>\mu_0$ **有相同的拒绝域.**



证明 在检验问题 $H_0:\mu\le\mu_0$, $H_1:\mu>\mu_0$ 中,

因为 H_0 中的 μ 都比 H_1 中的 μ 小,

从直观上看, 合理的检验法则是:

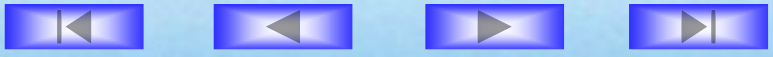
若观察值 $\bar{x}$ 与 μ_0 的差 $\bar{x}-\mu_0$ 过分大, 即 $\bar{x}-\mu_0\ge k$,

则我们拒绝 H_0 接受 H_1.

拒绝域的形式 $\bar{x}-\mu_0\ge k$, (k 待定).

由标准正态分布的分布函数 $\Phi(x)$ 的单调性可知,

$$P\{\text{拒绝 } H_0 \mid H_0 \text{ 为真}\}=P_{\mu\le\mu_0}(\bar{x}-\mu_0\ge k)$$



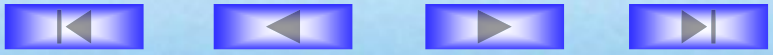
$$= P_{\mu\le\mu_0}\left(\frac{\bar{x}-\mu_0}{\sigma/\sqrt{n}} \ge \frac{\mu_0+k-\mu}{\sigma/\sqrt{n}}\right)$$

$$= 1-\Phi\left(\frac{(\mu_0+k)-\mu}{\sigma/\sqrt{n}}\right)_{\mu\le\mu_0} = \Phi\left(\frac{\mu-(\mu_0+k)}{\sigma/\sqrt{n}}\right)_{\mu\le\mu_0}$$

$$\le \Phi\left(\frac{\mu_0-(\mu_0+k)}{\sigma/\sqrt{n}}\right) = \Phi\left(\frac{-k}{\sigma/\sqrt{n}}\right),$$

因此要控制$P\{$拒绝 $H_0 \mid H_0$ 为真$\} \le \alpha$,

只需令 $\Phi\left(\dfrac{-k}{\sigma/\sqrt{n}}\right) \le \alpha$, 即 $k = (\sigma/\sqrt{n})z_\alpha$,



检验问题 $H_0:\mu\le\mu_0$, $H_1:\mu>\mu_0$ 的拒绝域为

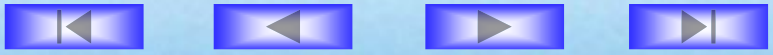
$$\bar{x}-\mu_0\ge(\sigma/\sqrt{n})z_\alpha,\quad \text{即}\ \frac{\bar{x}-\mu_0}{\sigma/\sqrt{n}}\ge z_\alpha.$$

比较正态总体$N(\mu,\sigma^2)$在方差σ^2已知时, 对均值μ的两种检验问题

$H_0:\mu=\mu_0, H_1:\mu>\mu_0$ 和 $H_0:\mu\le\mu_0, H_1:\mu>\mu_0$,

尽管原假设H_0的形式不同,实际意义也不同, 但对于相同的显著性水平α, 它们的拒绝域相同

第二类形式的检验问题可归结为第一类形式讨论.



2. σ^2为未知, 关于μ的检验(t检验)

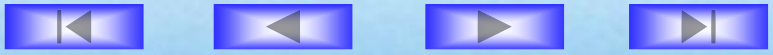
设总体$X \sim N(\mu,\sigma^2)$, 其中μ,σ^2未知, **我们来求检验问题**

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

的拒绝域 (显著性水平为α).

设 $X_1,X_2,\cdots,X_n$ 为来自总体 X 的样本, **由于** σ^2 未知, 现在不能利用$\dfrac{\overline{X}-\mu_0}{\sigma/\sqrt{n}}$来确定拒绝域 .

注意到S^2 是 σ^2 的无偏估计, 我们用S 来代替σ,



采用

$$t=\frac{\overline{X}-\mu_0}{S/\sqrt{n}}$$

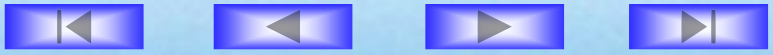
来作为检验统计量. 当观察值 $|t|=\left|\frac{\overline{x}-\mu_0}{s/\sqrt{n}}\right|$ 过分大时就拒绝 H_0, 拒绝域的形式为

$$|t|=\left|\frac{\overline{x}-\mu_0}{s/\sqrt{n}}\right|\geq k.$$

根据第六章§2定理三

定理三

当H_0为真时, $\frac{\overline{X}-\mu_0}{S/\sqrt{n}}\sim t(n-1)$, 故由



$$P\{\text{当 } H_0 \text{ 为真, 拒绝 } H_0\} = P_{\mu_0}\left\{\left|\frac{\overline{X}-\mu_0}{S/\sqrt{n}}\right| \geq k\right\} = \alpha,$$

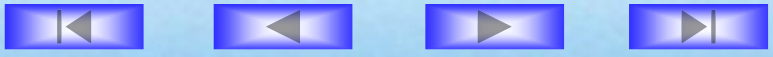
得 $k = t_{\alpha/2}(n-1)$, 即得

拒绝域为 $|t| = \left|\dfrac{\overline{x}-\mu_0}{s/\sqrt{n}}\right| \geq t_{\alpha/2}(n-1)$.

对于正态总体 $N(\mu,\sigma^2)$, 当 σ^2 未知时, 关于 μ 的单边检验的拒绝域在表 8.1 中给出.

上述利用 t 统计量得出的检验法称为 t 检验法.

在实际中, 正态总体的方差常为未知, 所以我们常用 t 检验法来检验关于正态总体均值的检验问题.



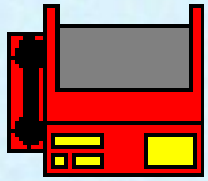
例1　某种电子元件的寿命X(以小时计)服从正态分布，μ,σ^2 均为未知. 现测得16只元件的寿命如下:

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(小时)?

解　依题意需检验假设

$$H_0:\mu\le\mu_0=225,\quad H_1:\mu>225,$$



取 $\alpha=0.05,\quad n=16,\quad \bar{x}=241.5,\quad s=98.7259,$

t表分布

查表得

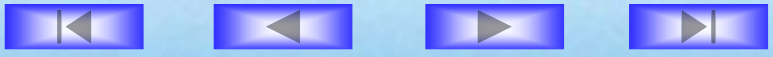
$$t_{0.05}(15)=1.7531 > \left|t\right|=\left|\frac{\bar{x}-\mu_0}{s/\sqrt{n}}\right|=0.6685$$

故接受H_0, 认为元件的平均寿命不大于225小时.

二、两个总体$N(\mu_1,\sigma_1^2)$, $N(\mu_2,\sigma_2^2)$的情况

利用t检验法检验具有相同方差的两正态总体均值差的假设.

设 $X_1,X_2,\cdots,X_{n_1}$ 为来自正态总体 $N(\mu_1,\sigma^2)$ 的样本, $Y_1,Y_2,\cdots,Y_{n_2}$ 为来自正态总体 $N(\mu_2,\sigma^2)$的样本,且设两样本独立 .又设 $\overline{X},\overline{Y}$ 分别是总体的样本均值, S_1^2,S_2^2是样本方差,μ_1,μ_2,σ^2均为未知,要特别注意的是,这里假设两总体的方差相等.



现在来求检验问题 ：

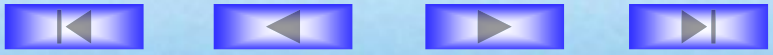
$$H_0:\mu_1-\mu_2=\delta,\quad H_1:\mu_1-\mu_2\neq\delta$$

(δ 为已知常数)的拒绝域 . 取显著性水平为α.

引入下述t 统计量作为检验统计量 ：

$$t=\frac{(\overline{X}-\overline{Y})-\delta}{S_w\sqrt{\frac{1}{n_1}+\frac{1}{n_2}}},$$

其中 $S_w^2=\dfrac{(n_1-1)S_1^{\,2}+(n_2-1)S_2^{\,2}}{n_1+n_2-2}$, $S_w=\sqrt{S_w^2}$.



由 $P\{H_0\text{为真，拒绝}H_0\}$

$$= P_{\mu_1-\mu_2=\delta}\left\{\frac{|(\bar{X}-\bar{Y})-\delta|}{S_w\sqrt{\frac{1}{n_1}+\frac{1}{n_2}}}\geq k\right\}=\alpha$$

可得 $k=t_{\alpha/2}(n_1+n_2-2)$. 故得拒绝域为

$$t=\frac{|(\bar{x}-\bar{y})-\delta|}{s_w\sqrt{\frac{1}{n_1}+\frac{1}{n_2}}}\geq t_{\alpha/2}(n_1+n_2-2).$$

关于均值差的其他两个检验问题的拒绝域见表8.1，常用的是$\delta=0$的情况.

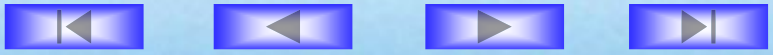
当两个正态总体的方差均为已知(不一定相等) 时, 我们可用 Z 检验法来检验两正态总体均值差的假设问题, 见表8.1 .

例2 用两种方法（A和B）测定冰自−0.72°C转变为0°C的水的融化热(以卡 / 克计). 测得以下数据：

方法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



方法B：

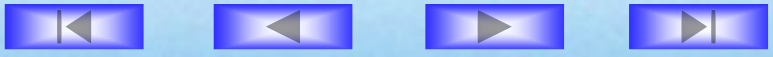
80.02 79.94 79.98 79.97 79.97 80.03 79.95 78.97

设这两个样本相互独立，且分别来自正态总体 $N(\mu_1,\sigma^2)$和$N(\mu_2,\sigma^2)$,μ_1,μ_2,σ^2均未知.试检验假设（取显著性水平$\alpha=0.05$)：

$$H_0=\mu_1-\mu_2\le 0,$$

$$H_1=\mu_1-\mu_2>0,$$

解　分别画出对应于方法A和方法B的数据的箱线图(图8－3略)，如图这两种方法所得的结果具有明显差异的，现在来检验上述我们看到的假设.



$$n_1 = 13,\quad \bar{x}_A = 80.02,$$

$$s_A^2 = 0.024^2,$$

$$n_2 = 8,\quad \bar{x}_B = 79.98,$$

$$s_B^2 = 0.03^2,$$

$$s_w^2 = \frac{12\times s_A^2 + 7\times s_B^2}{19} = 0.0007178.$$

$$t = \frac{\bar{x}_A - \bar{x}_B}{s_w\sqrt{1/13+1/8}} = 3.33 > t_{0.05}(13+8-2) = 1.7291.$$

故拒绝H_0, 认为方法 A 比方法 B 测得的融化热要大.

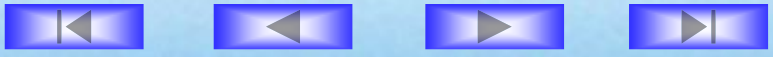
三、基于成对数据的检验(t 检验)

有时为了比较两种产品, 或两种仪器, 两种方法等的差异, 我们常在相同的条件下做对比试验, 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3　有两台光谱仪 I_x，I_y，用来测量材料中某种金属的含量，为鉴定它们的测量结果有无显著差异，制备了9件试块(它们的成分、金属含量、均匀性等各不相同)，现在分别用这两台机器对每一试块测量一次，得到9对观察值如下：

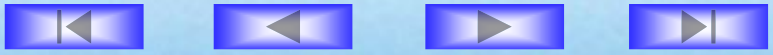
$x(\%)$	0.20	0.30	0.40	0.50	0.60	0.70	0.80	0.90	1.00
$y(\%)$	0.10	0.21	0.52	0.32	0.78	0.59	0.68	0.77	0.89
$d=x-y(\%)$	0.10	0.09	−0.12	0.18	−0.18	0.11	0.12	0.13	0.11

问能否认为这两台仪器的测量结果有显著的差异？
$(\alpha=0.01)$



解　本题中的数据是成对的, 即对同一试块一对数据, 我们看到一对与另一对之间的差异是由各种因素, 如材料成分、金属含量、均匀性等因素引起的. 由于各试块的特性有广泛的差别, 表中第一行不能看成是一个样本的样本值. 表中第二行也不能看成是一个样本的样本值.

而同一对中两个数据的差异则可看成是仅由这两台仪器性能的差异所引起的. 这样, 局限于各对中两个数据来比较就能排除种种其他因素, 而只考虑单独由仪器的性能所产生的影响.



表中第三行表示各对数据的差 $d_i = x_i - y_i$,

设 $D_i = X_i - Y_i (i = 1,2,\cdots,n)$来自正态总体$N(\mu_D, \sigma_D^2)$,

这里μ_D, σ_D^2均为未知. **若两台机器的性能一样,**

则各对数据的差异$d_1, d_2, \cdots, d_n$ 属随机误差,

随机误差可以认为服从正态分布, 其均值为零.

需检验假设

$$H_0 : \mu_D = 0, \quad H_1 : \mu_D \neq 0;$$

设 $D_1, D_2, \cdots, D_n$ 的样本均值 $\bar{d}$, 样本方差 s_D^2,

拒绝域为 $|t| = \left|\frac{\bar{d} - 0}{s_D / \sqrt{n}}\right| \geq t_{\alpha/2}(n-1)$,

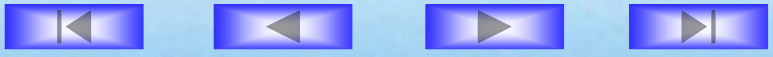
现在 $n = 9,\ t_{\alpha/n} = t_{0.005}(8) = 3.3554$, 即知拒绝域为

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

由观察值得 $\bar{d} = 0.06,\ s_D = 0.1227,$

$$t = \frac{0.06}{0.1227/\sqrt{9}} = 1.467 < 3.3554$$

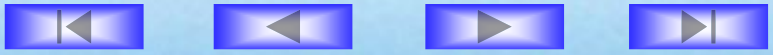
现$|t|$的值不落在拒绝域内,故接受H_0,认为两台机器的测量成果并无显著差异.



例4　做以下的实验以比较人对红光或绿光的反应时间（以秒计）．实验在点亮红光或绿光的同时，启动计时器，要求受试者见到红光或绿光点亮时，就按下按钮，切断记时器，这就能测得反应时间. 测得的结果如下表：

红光(x)	0.30	0.23	0.41	0.53	0.24	0.36	0.38	0.51
绿光(y)	0.43	0.32	0.58	0.46	0.27	0.41	0.38	0.61
$d=x-y$	−0.13	−0.09	−0.17	0.07	−0.03	−0.05	0.00	−0.10

设$D_i = X_i - Y_i$ $(i = 1,2,\cdots,8)$ 是来自正态总体$N(\mu_D,\sigma_D^2)$的样本，μ_D,σ_D^2均未知. 试检验假设(取



显著性水平$\alpha = 0.05$)

$$H_0 : \mu_D \geq 0, \quad H_1 : \mu_D < 0;$$

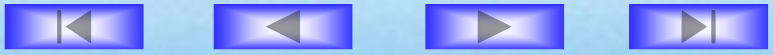
解 现在$n = 8$, $\overline{x_d} = -0.0625$, $s_d = 0.0765$,

而

$$\frac{\overline{x_d}}{s_d / \sqrt{8}} = -2.311 < -t_{0.05}(7) = -1.8946$$

故拒绝 H_0, 认为$\mu_D < 0$, 即认为人对红光的反应时间小于对绿光的反应时间，**也就是人对红光的反应要比绿光快.**

补充例题



四、小结

本节学习的正态总体均值的假设检验有:

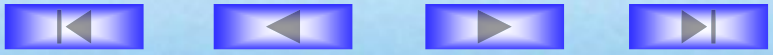
1. 单个总体均值 μ 的检验 — Z 检验;

2. 两个总体均值差 $\mu_1 - \mu_2$ 的检验 — t 检验;

3. 基于成对数据的检验 — t 检验;

正态总体均值、方差的检验法见下表

(显著性水平为 α)

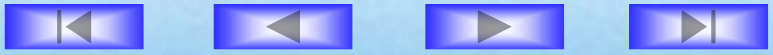


假设检验的基本原理、相关概念和一般步骤.

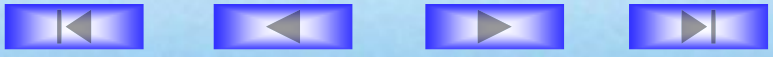
假设检验的两类错误

真实情况(未知)	所　作　决　策	
	接受 H_0	拒绝 H_0
H_0 为真	正确	犯第I类错误
H_0 不真	犯第II类错误	正确

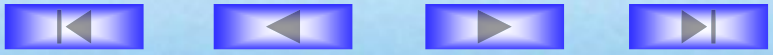
课后作业：教材262页习题第1，2，3题



	原假设H_0	检验统计量	备择假设H_1	拒绝域
1	$\mu \le \mu_0$ $\mu \ge \mu_0$ $\mu = \mu_0$ (σ^2已知)	$Z = \dfrac{\overline{X} - \mu_0}{\sigma / \sqrt{n}}$	$\mu > \mu_0$ $\mu < \mu_0$ $\mu \ne \mu_0$	$z \ge z_\alpha$ $z \le -z_\alpha$ $\lvert z \rvert \ge z_{\alpha/2}$
2	$\mu \le \mu_0$ $\mu \ge \mu_0$ $\mu = \mu_0$ (σ^2未知)	$t = \dfrac{\overline{X} - \mu_0}{S / \sqrt{n}}$	$\mu > \mu_0$ $\mu < \mu_0$ $\mu \ne \mu_0$	$t \ge t_\alpha(n-1)$ $t \le -t_\alpha(n-1)$ $\lvert t \rvert \ge t_{\alpha/2}(n-1)$
3	$\mu_1 - \mu_2 \le \delta$ $\mu_1 - \mu_2 \ge \delta$ $\mu_1 - \mu_2 = \delta$ (σ_1^2, σ_2^2已知)	$Z = \dfrac{\overline{X} - \overline{Y} - \delta}{\sqrt{\dfrac{\sigma_1^2}{n_1} + \dfrac{\sigma_2^2}{n_2}}}$	$\mu - \mu_0 > \delta$ $\mu - \mu_0 < \delta$ $\mu - \mu_0 \ne \delta$	$z \ge z_\alpha$ $z \le -z_\alpha$ $\lvert z \rvert \ge z_{\alpha/2}$
4	$\mu_1 - \mu_2 \le \delta$ $\mu_1 - \mu_2 \ge \delta$ $\mu_1 - \mu_2 = \delta$ ($\sigma_1^2 = \sigma_2^2 = \sigma^2$未知)	$t = \dfrac{\overline{X} - \overline{Y} - \delta}{S_w \sqrt{\dfrac{1}{n_1} + \dfrac{1}{n_2}}}$ $S_w^2 = \dfrac{(n_1-1)S_1^2 + (n_2-2)S_2^2}{n_1 + n_2 - 2}$	$\mu - \mu_0 > \delta$ $\mu - \mu_0 < \delta$ $\mu - \mu_0 \ne \delta$	$t \ge t_\alpha(n_1 + n_2 - 2)$ $t \le -t_\alpha(n_1 + n_2 - 2)$ $\lvert t \rvert \ge t_{\alpha/2}(n_1 + n_2 - 1)$



	原假设H_0	检验统计量	备择假设H_1	拒绝域
5	$\sigma^2 \le \sigma_0^2$ $\sigma^2 \ge \sigma_0^2$ $\sigma^2 = \sigma_0^2$ (μ未知)	$\chi^2 = \dfrac{(n-1)S^2}{\sigma_0^2}$	$\sigma^2 > \sigma_0^2$ $\sigma^2 < \sigma_0^2$ $\sigma^2 \ne \sigma_0^2$	$\chi^2 \ge \chi_\alpha^2(n-1)$ $\chi^2 \le \chi_{1-\alpha}^2(n-1)$ $\chi^2 \ge \chi_{\alpha/2}^2(n-1)$或 $\chi^2 \le \chi_{1-\alpha/2}^2(n-1)$
6	$\sigma_1^2 \le \sigma_2^2$ $\sigma_1^2 \ge \sigma_2^2$ $\sigma_1^2 = \sigma_2^2$ (μ_1, μ_2未知)	$F = \dfrac{S_1^2}{S_2^2}$	$\sigma_1^2 > \sigma_2^2$ $\sigma_1^2 < \sigma_2^2$ $\sigma_1^2 \ne \sigma_2^2$	$F \ge F_\alpha(n_1-1, n_2-1)$ $F \le F_{1-\alpha}(n_1-1, n_2-1)$ $F \ge F_{\alpha/2}(n_1-1, n_2-1)$或 $F \ge F_{1-\alpha/2}(n_1-1, n_2-1)$
7	$\mu_D \le 0$ $\mu_D \ge 0$ $\mu_D = 0$ (成对数据)	$t = \dfrac{\overline{D}-0}{S_D/\sqrt{n}}$	$\mu_D > 0$ $\mu_D < 0$ $\mu_D \ne 0$	$t \ge t_\alpha(n-1)$ $t \le -t_\alpha(n-1)$ $\lvert t \rvert \ge t_{\alpha/2}(n-1)$



第三节 正态总体方差的假设检验

一、单个正态总体方差的假设检验

二、两个正态总体方差的假设检验

三、小结

一、单个正态总体方差的假设检验

设总体 $X \sim N(\mu,\sigma^2)$, μ,σ^2均为未知,

$X_1, X_2,\cdots, X_n$ 为来自总体 X 的样本.

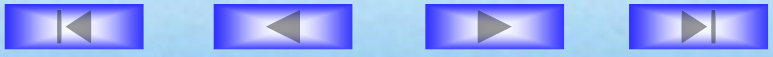
要求检验假设: (显著性水平为α)

$$H_0:\sigma^2=\sigma_0^{\,2},\quad H_1:\sigma^2\neq\sigma_0^{\,2},$$

σ_0 为已知常数 .

由于 S^2 是 σ^2 的无偏估计, 当H_0 为真时, 比值 $\dfrac{S^2}{\sigma_0^{\,2}}$

在1附近摆动, 不应过分大于1或过分小于1,



根据第六章§3定理二可知， 当 H_0 为真时，

$$\frac{(n-1)S^2}{\sigma_0^2} \sim \chi^2(n-1),$$

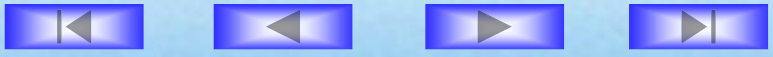
我们取

$$\chi^2 = \frac{(n-1)S^2}{\sigma_0^2}$$

作为检验统计量，**上述检验问题的拒绝域具有以下的形式：**

$$\frac{(n-1)S^2}{\sigma_0^2} \le k_1 \text{或} \frac{(n-1)S^2}{\sigma_0^2} \ge k_2,$$

此处 k_1 和 k_2 的值由下式确定：



$$P\{H_0 \text{为真, 拒绝 } H_0\}$$

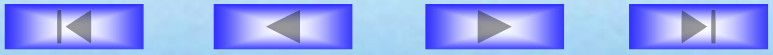
$$=P_{\sigma_0^2}\left\{\left(\frac{(n-1)S^2}{\sigma_0^2}\le k_1\right)\cup\left(\frac{(n-1)S^2}{\sigma_0^2}\ge k_2\right)\right\}=\alpha.$$

为了计算方便, 习惯上取

$$P_{\sigma_0^2}\left\{\frac{(n-1)S^2}{\sigma_0^2}\le k_1\right\}=\frac{\alpha}{2},\ p_{\sigma_0^2}\left\{\frac{(n-1)S^2}{\sigma_0^2}\ge k_2\right\}=\frac{\alpha}{2}$$

故得 $k_1=\chi^2_{1-\alpha/2}(n-1),\ \ k_2=\chi^2_{\alpha/2}(n-1).$

于是得拒绝域为:



$$\frac{(n-1)s^2}{\sigma_0^2} \le \chi^2_{1-\alpha/2}(n-1) \text{ 或 } \frac{(n-1)s^2}{\sigma_0^2} \ge \chi^2_{\alpha/2}(n-1).$$

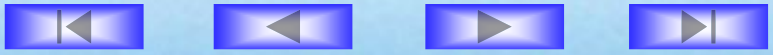
下面来求单边检验问题的拒绝域 (设显著水平为α)

$$H_0: \sigma^2 \le \sigma_0^2, \quad H_1: \sigma^2 > \sigma_0^2,$$

的拒绝域. 因H_0中的全部σ^2都比H_1中的σ^2要小，当H_1为真时，S^2的观察值s^2往往偏大,

因此拒绝域的形式为: $s^2 \ge k$.

此处k的值由下式确定:



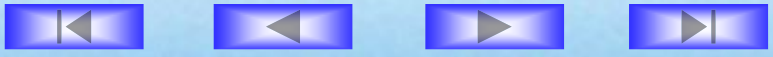
$$P\{H_0\text{为真，拒绝}H_0\}=P_{\sigma^2\le\sigma_0^2}\{S^2\ge k\}$$

$$=P_{\sigma^2\le\sigma_0^2}\left\{\frac{(n-1)S^2}{\sigma_0^2}\ge\frac{(n-1)k}{\sigma_0^2}\right\}$$

$$\le P_{\sigma^2\le\sigma_0^2}\left\{\frac{(n-1)S^2}{\sigma^2}\ge\frac{(n-1)k}{\sigma_0^2}\right\}.\quad(\text{因为}\sigma^2\le\sigma_0^2)$$

要使 $P\{H_0\text{为真，拒绝}H_0\}\le\alpha$，只需令

$$P_{\sigma^2\le\sigma_0^2}\left\{\frac{(n-1)S^2}{\sigma^2}\ge\frac{(n-1)k}{\sigma_0^2}\right\}=\alpha.$$



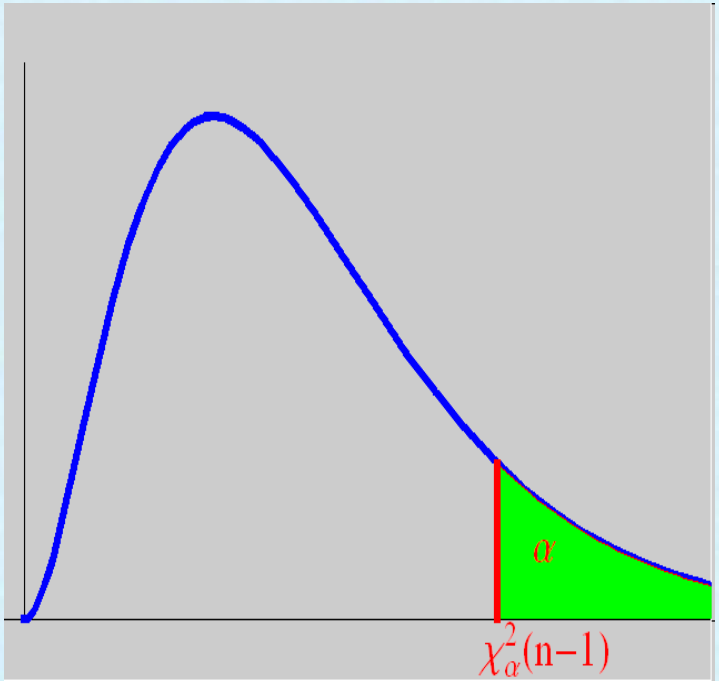
因 $\dfrac{(n-1)S^2}{\sigma^2} \sim \chi^2(n-1)$, $\dfrac{(n-1)k}{\sigma_0^{\,2}} = \chi_\alpha^2(n-1)$

于是 $k = \dfrac{\sigma_0^{\,2}}{n-1}\chi_\alpha^2(n-1)$,

右边检验问题的拒绝域为

$$s^2 \geq \frac{\sigma_0^{\,2}}{n-1}\chi_\alpha^2(n-1),$$

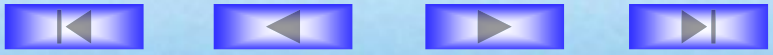
$$\text{即 } \chi^2 = \frac{(n-1)s^2}{\sigma_0^{\,2}} \geq \chi_\alpha^2(n-1).$$



同理左边检验问题: $H_0 : \sigma^2 \geq \sigma_0^{\,2}$, $H_1 : \sigma^2 < \sigma_0^{\,2}$,

拒绝域为 $$\chi^2=\frac{(n-1)s^2}{\sigma_0^{\,2}}\le\chi^2_{1-\alpha}(n-1).$$

上述检验法称为χ^2 检验法.



例1 某厂生产的某种型号的电池，其寿命长期以来服从方差$\sigma^2 = 5000$的正态分布，现有一批这种电池，从它生产情况来看，寿命的波动性有所变化. 现随机的取26只电池，测出其寿命的样本方差$s^2 = 9200$，问根据这一数据能否推断这批电池的寿命的波动性较以往的有显著的变化？

(取$\alpha = 0.02$)

解 本题要求在水平$\alpha = 0.02$下检验假设

$$H_0 : \sigma^2 = 5000, \quad H_1 : \sigma^2 \neq 5000.$$

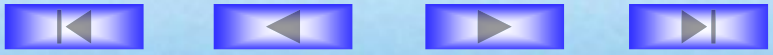
现在 $\alpha = 0.02,\ n = 26,$

$$\chi^2_{\alpha/2}(n-1) = \chi^2_{0.01}(25) = 44.31,$$

$$\chi^2_{1-\alpha/2}(n-1) = \chi^2_{0.99}(25) = 11.52,$$

$\sigma_0^{\,2} = 5000,$ **由(3.1)拒绝域为**

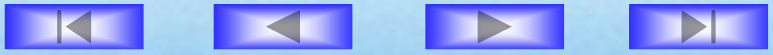
$$\frac{(n-1)s^2}{\sigma_0^{\,2}} \leq 11.52, \text{ 或 } \frac{(n-1)s^2}{\sigma_0^{\,2}} \geq 44.31.$$



由观察值$s^2 = 9200$得

$$\frac{(n-1)s^2}{\sigma_0^2} = \frac{25 \times 9200}{5000} = 46 > 44.314,$$

所以拒绝H_0，**认为这批电池的寿命的波动性较以往的有显著的变化.**

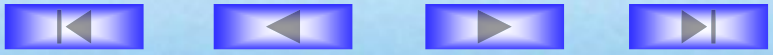


二、两个正态总体方差的假设检验

设 $X_1, X_2, \cdots, X_{n_1}$ 为来自正态总体 $N(\mu_1, \sigma_1^2)$的样本，设 $Y_1, Y_2, \cdots, Y_{n_1}$ 为来自正态总体 $N(\mu_2, \sigma_2^2)$的样本，且设两样本独立，其样本方差为 S_1^2, S_2^2. 又设 $\mu_1, \mu_2, \sigma_1^2, \sigma_2^2$均为未知，**现在需要检验假设:**（显著性水平为α）

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

当H_0为真时，$E(S_1^2) = \sigma_1^2 \le \sigma_2^2 = E(S_2^2)$,



当H_1为真时，$E(S_1^2)=\sigma_1^2>\sigma_2^2=E(S_2^2)$,

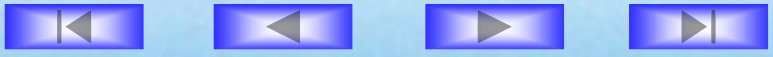
当H_1为真时，观察值$\dfrac{S_1^2}{S_2^2}$有偏大的趋势，

故拒绝域的形式为$\dfrac{S_1^2}{S_2^2}\geq k$, 常数$k$的值由下式确定：

$$P\{H_0\text{为真，拒绝}H_0\}=P_{\sigma_1^2\leq\sigma_2^2}\left\{\frac{S_1^2}{S_2^2}\geq k\right\}$$

$$\leq P_{\sigma_1^2\leq\sigma_2^2}\left\{\frac{S_1^2/S_2^2}{\sigma_1^2/\sigma_2^2}\geq k\right\},\ (\text{因为}\sigma_1^2/\sigma_2^2\leq 1)$$

要使 $P\{H_0\text{为真，拒绝}H_0\}\leq\alpha$, 只需令



$$P_{\sigma_1^2 \le \sigma_2^2}\left\{\frac{S_1^2/S_2^2}{\sigma_1^2/\sigma_2^2} \ge k\right\} = \alpha.$$

由第六章§3定理四知

定理四

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

$$k = F_\alpha(n_1-1,\ n_2-1).$$

即得检验问题的拒绝域为

$$F = \frac{S_1^2}{S_2^2} \ge F_\alpha(n_1-1,\ n_2-1).$$

上述检验法称为

***F* 检验法.**

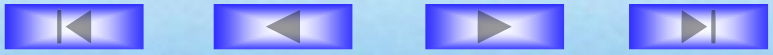
例2 设第2节例 2中的两个样本分别来自总体 $N(\mu_A,\sigma_A^2), N(\mu_B,\sigma_B^2)$, 且两样本独立. **试检验** $H_0:\sigma_A^2=\sigma_B^2$, $H_1:\sigma_A^2\neq\sigma_B^2$, 以说明我们假设 $\sigma_A^2=\sigma_B^2$是合理的.(取显著性水平$\alpha=0.01$.)

解 此处$n_1=13$, $n_2=8$, $\alpha=0.01$, **拒绝域为**

$$\frac{S_A^2}{S_B^2}\geq F_{0.005}(12,7)=8.18,$$

或

$$\frac{S_A^2}{S_B^2}\leq F_{0.005}(12,7)=\frac{1}{F_{0.005}(7,12)}=\frac{1}{5.52}=0.18.$$



现在$s_A^2=(0.024)^2$，$s_B^2=(0.03)^2$，$s_A^2/s_B^2=0.64$,

$$0.18<0.64<8.18,$$

故接受H_0,认为两总体方差相等.两总体方差相等

也称两总体具有方差齐性，这也表明第2节例2

假设两总体方差相等是合理的.

补充例题

三、小结

1. 单个正态总体方差的检验法—χ^2 检验法;

2. 两个正态总体方差的检验法—F 检验法;

正态总体均值、方差的检验法见下表

(显著性水平为α)

第四节 置信区间与假设检验之间的关系

一、置信区间与双边检验之间的对应关系

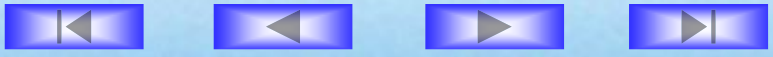
二、 置信区间与单边检验之间的对应关系

三、小结

一、置信区间与双边检验之间的对应关系

设 $X_1, X_2, \cdots, X_n$ 是一个来自总体的样本，$x_1, x_2, \cdots, x_n$ 是相应的样本值，Θ 是参数 θ 的可能取值范围.设 $(\underline{\theta}(X_1, X_2, \cdots, X_n), \overline{\theta}(X_1, X_2, \cdots, X_n))$ 是参数θ的一个置信水平为$1-\alpha$ 的置信区间，则对于任意的$\theta \in \Theta$, 有

$$P_\theta\{\underline{\theta}(X_1, X_2, \cdots, X_n) < \theta < \overline{\theta}(X_1, X_2, \cdots, X_n)\} \geq 1-\alpha,$$



考虑显著水平为 α 的双边检验：

$$H_0:\theta=\theta_0,\quad H_1:\theta\neq\theta_0.$$

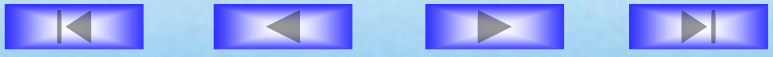
因为

$$P_{\theta_0}\{\underline{\theta}(X_1,X_2,\cdots,X_n)<\theta_0<\overline{\theta}(X_1,X_2,\cdots,X_n)\}$$
$$\geq 1-\alpha,$$

即有

$$P_{\theta_0}\{(\theta_0\leq\underline{\theta}(X_1,X_2,\cdots,X_n))\cup(\theta_0\geq\overline{\theta}(X_1,X_2,\cdots,X_n))\}$$
$$\leq\alpha.$$

按显著性水平为α的假设检验的拒绝域的定义，



$\theta_0 \le \underline{\theta}(x_1, x_2, \cdots, x_n)$ 或 $\theta_0 \ge \overline{\theta}(x_1, x_2, \cdots, x_n)$;

接受域为

$$\underline{\theta}(x_1, x_2, \cdots, x_n) < \theta_0 < \overline{\theta}(x_1, x_2, \cdots, x_n).$$

当我们要检验假设 $H_0 : \theta = \theta_0$, $H_1 : \theta \ne \theta_0$时,

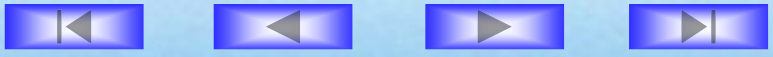
先求出θ的置信水平为 $1-\alpha$ 的置信区间 $(\underline{\theta}, \overline{\theta})$,

若 $\theta_0 \in (\underline{\theta}, \overline{\theta})$, 则接受 H_0;

若 $\theta_0 \notin (\underline{\theta}, \overline{\theta})$, 则拒绝 H_0.

反之,对于任意的 $\theta_0 \in \Theta$,

考虑显著性水平为α的假设检验问题:



$$H_0:\theta=\theta_0,\qquad H_1:\theta\neq\theta_0.$$

假设它的接受域为

$$\underline{\theta}(x_1,x_2,\cdots,x_n)<\theta_0<\overline{\theta}(x_1,x_2,\cdots,x_n).$$

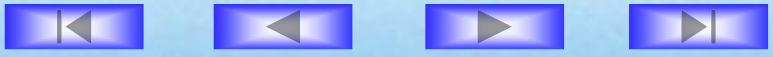
即有

$$P_{\theta_0}\{\underline{\theta}(X_1,X_2,\cdots,X_n)<\theta_0<\overline{\theta}(X_1,X_2,\cdots,X_n)\}$$

由θ_0的任意性, 有

$$P_{\theta}\{\underline{\theta}(X_1,X_2,\cdots,X_n)<\theta<\overline{\theta}(X_1,X_2,\cdots,X_n)\}$$

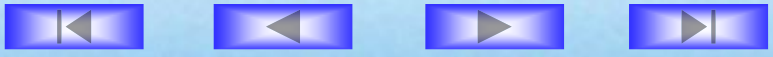
因此$(\underline{\theta}(X_1,X_2,\cdots,X_n),\ \overline{\theta}(X_1,X_2,\cdots,X_n))$是参数$\theta$的一个置信水平为$1-\alpha$的置信区间.



这就是说，为要求出参数 θ 的置信水平为 $1-\alpha$ 的置信区间，要先求出显著水平为 α 的检验假设 $H_0:\theta=\theta_0$，$H_1:\theta\neq\theta_0$，的接受域：

$$\underline{\theta}(x_1,x_2,\cdots,x_n)<\theta_0<\overline{\theta}(x_1,x_2,\cdots,x_n).$$

那么，$(\underline{\theta}(X_1,X_2,\cdots,X_n),\overline{\theta}(X_1,X_2,\cdots,X_n))$ 是参数 θ 的一个置信水平为 $1-\alpha$ 的置信区间 .



二、置信区间与单边检验之间的对应关系

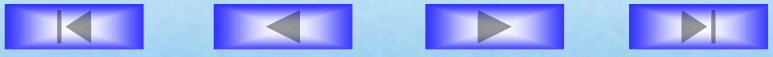
(1) 置信水平为$1-\alpha$的单侧置信区间

$(-\infty, \bar{\theta}(X_1, X_2, \cdots, X_n))$与显著水平为$\alpha$的左边检验问题 $H_0: \theta \geq \theta_0$, $H_1: \theta < \theta_0$ 有类似的对应关系

若已求得单侧置信区间 $(-\infty, \bar{\theta}(X_1, X_2, \cdots, X_n))$,

则当$\theta_0 \in (-\infty, \bar{\theta}(x_1, x_2, \cdots, x_n))$时接受$H_0$;

当$\theta_0 \notin (-\infty, \bar{\theta}(x_1, x_2, \cdots, x_n))$时拒绝$H_0$.



反之, 若已求得检验问题 $H_0:\theta\geq\theta_0, H_1:\theta<\theta_0$ 的接受域为:

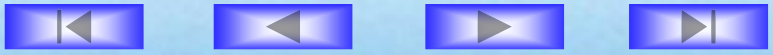
$$-\infty<\theta_0\leq\overline{\theta}(x_1,x_2,\cdots,x_n),$$

则可得θ的一个单侧置信区间

$$(-\infty,\overline{\theta}(X_1,X_2,\cdots,X_n)).$$

(2) 置信水平为$1-\alpha$的单侧置信区间

$(\underline{\theta}(X_1,X_2,\cdots,X_n),+\infty)$与显著水平为$\alpha$的右边检验问题:



$H_0:\theta\le\theta_0$，$H_1:\theta>\theta_0$ 也有类似的对应关系 .

若已求得单侧置信区间 $(\underline{\theta}(X_1,X_2,\cdots,X_n),+\infty)$,

则当 $\theta_0\in(\underline{\theta}(x_1,x_2,\cdots,x_n),+\infty)$ 时接受 H_0;

当 $\theta_0\notin(\overline{\theta}(x_1,x_2,\cdots,x_n),+\infty)$ 时拒绝 H_0 .

反之, 若已求得检验问题 $H_0:\theta\le\theta_0$，$H_1:\theta>\theta_0$

的接受域为:

$$\underline{\theta}(x_1,x_2,\cdots,x_n)\le\theta_0<+\infty,$$

则可得 θ 的一个单侧置信区间

$$(\underline{\theta}(X_1,X_2,\cdots,X_n),+\infty).$$

例1 设 $X \sim N(\mu, 1)$, μ未知, $\alpha = 0.05$, $n = 16$,

且由一样本算得 $\bar{x} = 5.20$,

于是得到参数μ的一个置信水平为0.95的置信

区间 $\left(\bar{x} - \frac{1}{\sqrt{16}} z_{0.025},\ \bar{x} + \frac{1}{\sqrt{16}} z_{0.025}\right)$

$= (5.20 - 0.49,\ 5.20 + 0.49)$

$= (4.71,\ 5.69).$

考虑检验问题 $H_0 : \mu = 5.5$, $H_1 : \mu \neq 5.5$,

因为 $5.5 \in (4.71,\ 5.69)$, 所以接受H_0.

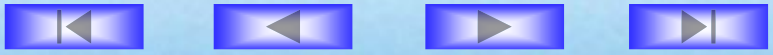
例2 数据如上例. 试求右边检验问题

$H_0:\mu\le\mu_0$, $H_1:\mu>\mu_0$ 的接受域, 并求 μ 的单侧置信下限. $(\alpha=0.05)$

解 **检验问题的拒绝域为** $z=\dfrac{\bar{x}-\mu_0}{\dfrac{1}{\sqrt{16}}}\ge z_{0.05}$,

即 $\mu_0\le 4.79$, **故检验问题的接受域为** $\mu_0>4.79$,

单侧置信区间$(4.79,+\infty)$, 单侧置信下限 $\underline{\mu}=4.79$.



三、小结

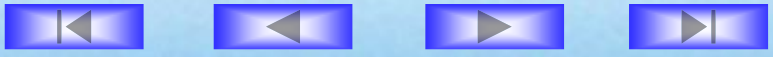
1. 置信区间与双边检验

$(\underline{\theta}(X_1, X_2, \cdots, X_n), \overline{\theta}(X_1, X_2, \cdots, X_n))$ 是参数 θ 的一个置信水平为 $1-\alpha$ 的置信区间.

2. 置信区间与单边检验

左边检验 θ 的单侧置信区间 $(-\infty, \overline{\theta}(X_1, X_2, \cdots, X_n))$.

右边检验 θ 的单侧置信区间 $(\underline{\theta}(X_1, X_2, \cdots, X_n), +\infty)$.



第五节 样本容量的选取

一、施行特征函数

二、Z 检验法的OC 函数

三、 t 检验法的OC 函数

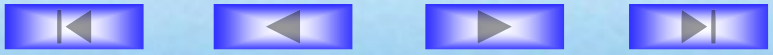
四、小结

一、施行特征函数

在一些实际问题中，我们除了希望控制犯第I类错误的概率外，往往还希望控制犯第 II类错误的概率.

以上在进行假设检验时，总是根据问题的需要，预先给出显著性水平以控制犯第I类错误的概率，而犯第 II类错误的概率则依赖于样本容量的选择.

在本节中，我们将阐明如何选取样本的容量使得犯第 II类错误的概率控制在预先给定的限度内，为此，引入施行特征函数.



施行特征函数的定义:

若C是参数θ的某检验问题的一个检验法,

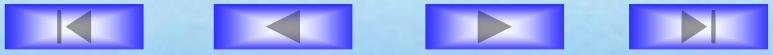
$$\beta(\theta) = P_0(\text{接受}H_0)$$

称为检验法C的施行特征函数或OC函数,

其图形称为OC曲线.

施行特征函数的作用:

适当地选取样本的容量, 使得犯第 II类错误的概率控制在预先给定的限度内.



二、Z 检验法的OC 函数

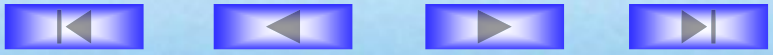
1. 右边检验问题

$H_0:\mu\le\mu_0, H_1:\mu>\mu_0$的$OC$ 函数是

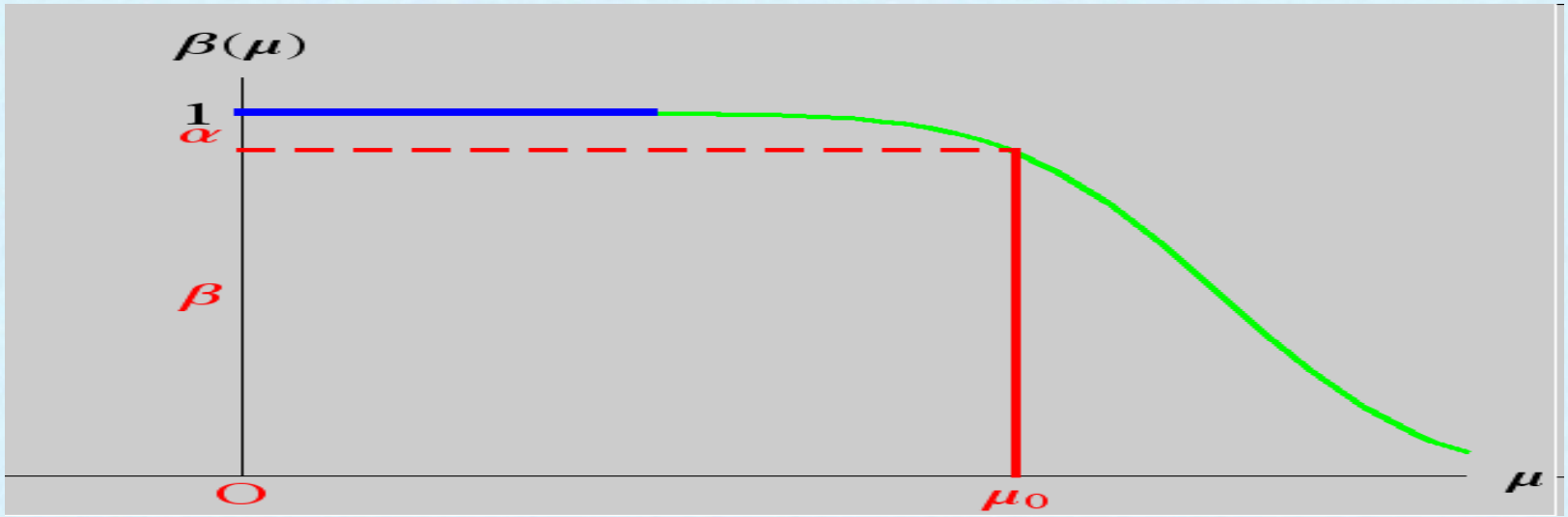
$$\beta(\mu)=P_\mu(\text{接受}H_0)=P_\mu\left\{\frac{\overline{X}-\mu_0}{\sigma/\sqrt{n}}<z_\alpha\right\}$$

$$=P_\mu\left\{\frac{\overline{X}-\mu}{\sigma/\sqrt{n}}<z_\alpha-\frac{\mu-\mu_0}{\sigma/\sqrt{n}}\right\}$$

$$=\Phi(z_\alpha-\lambda),\qquad \lambda=\frac{\mu-\mu_0}{\sigma/\sqrt{n}}.$$



此*OC* 函数的图形如下:



此*OC* 函数的性质如下:

(1)它是 $\lambda=\dfrac{\mu-\mu_0}{\sigma/\sqrt{n}}$ 的单调递减连续函数 ;

(2) $\lim\limits_{\mu\to\mu_0^+}\beta(\mu)=1-\alpha,\ \lim\limits_{\mu\to\infty}\beta(\mu)=0.$

根据 OC 函数 $\beta(\mu)$ 可以确定样本容量 n,
使当真值 $\mu \geq \mu_0 + \delta$ ($\delta > 0$为取定的值)时,
犯第II类错误的概率不超过给定的 β.

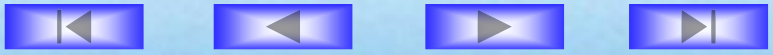
因为 $\beta(\mu)$ 是 μ 的递减函数,

故当 $\mu \geq \mu_0 + \delta$ 时, $\beta(\mu_0 + \delta) \geq \beta(\mu)$,

于是只要 $\beta(\mu_0 + \delta) = \Phi\left(z_\alpha - \dfrac{\sqrt{n}\delta}{\sigma}\right) \leq \beta$,

即 n 满足 $z_\alpha - \dfrac{\sqrt{n}\delta}{\sigma} \leq -z_\beta$, 只要 $\sqrt{n} \geq \dfrac{(z_\alpha + z_\beta)\sigma}{\delta}$,

就能使犯第 II类错误的概率不超过给定的 β.



2. 左边检验问题

假设检验 $H_0:\mu\geq\mu_0, H_1:\mu<\mu_0$ 的OC函数是

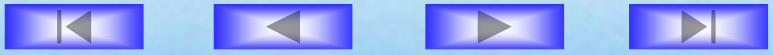
$$\beta(\mu)=P_\mu(\text{接受}H_0)=\Phi(z_\alpha+\lambda),\quad \lambda=\frac{\mu-\mu_0}{\sigma/\sqrt{n}}.$$

当真值$\mu\geq\mu_0$时$\beta(\mu)$为作出正确判断的概率;

当真值$\mu<\mu_0$时$\beta(\mu)$给出犯第II类错误的概率.

只要样本容量n满足 $\sqrt{n}\geq\dfrac{(z_\alpha+z_\beta)\sigma}{\delta}$

就能使犯第 II类错误的概率不超过给定的 β.



3. 双边检验问题

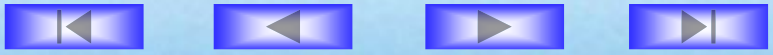
假设检验 $H_0:\mu=\mu_0, H_1:\mu\neq\mu_0$ 的OC函数是

$$\beta(\mu)=P_\mu(\text{接受}H_0)=P_\mu\left\{-z_{\alpha/2}<\frac{\overline{X}-\mu_0}{\sigma/\sqrt{n}}<z_{\alpha/2}\right\}$$

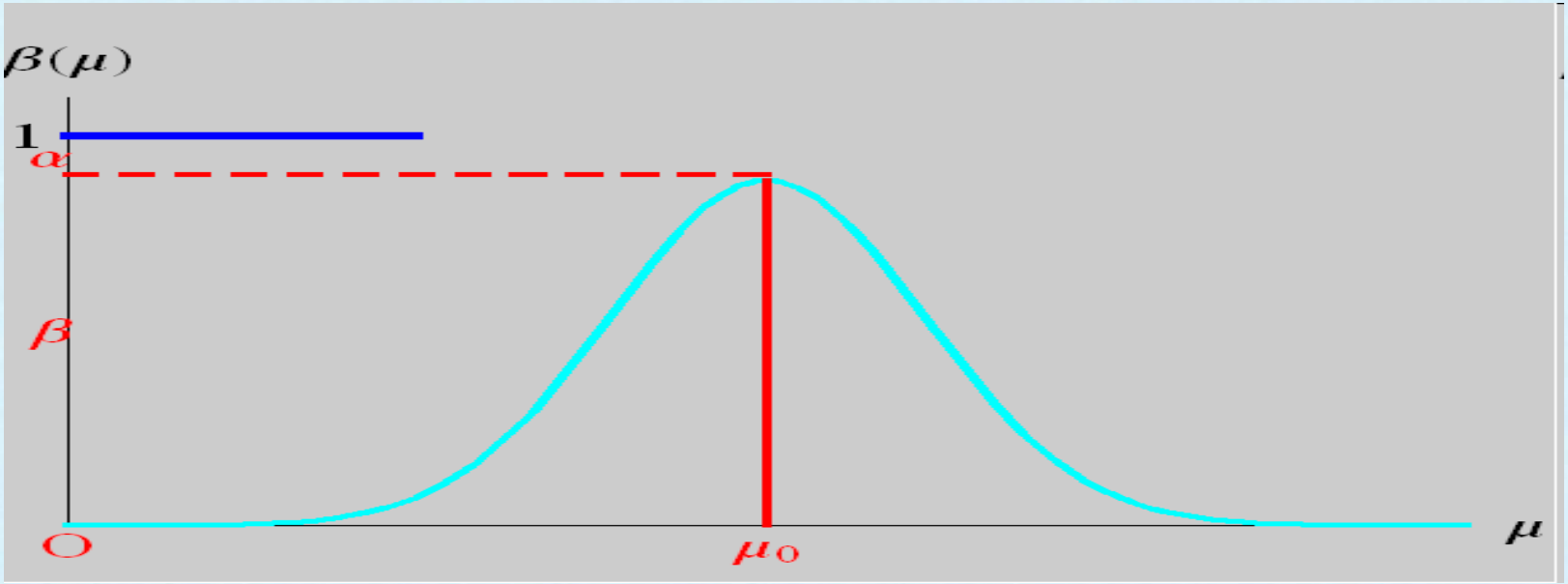
$$=P_\mu\left\{-\lambda-z_{\alpha/2}<\frac{\overline{X}-\mu}{\sigma/\sqrt{n}}<-\lambda+z_{\alpha/2}\right\}$$

$$=\Phi(z_{\alpha/2}-\lambda)-\Phi(-z_{\alpha/2}-\lambda)$$

$$=\Phi(z_{\alpha/2}-\lambda)+\Phi(z_{\alpha/2}+\lambda)-1,\quad \lambda=\frac{\mu-\mu_0}{\sigma/\sqrt{n}}.$$



此OC函数的图形如下:

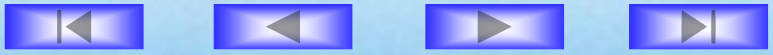


只要样本容量n满足 $\sqrt{n} \geq \dfrac{(z_{\alpha/2} + z_{\beta})\sigma}{\delta}$,

就能使犯第 II类错误的概率不超过给定的β.

例1　(工业产品质量抽验方案) 设有一大批产品, 产品质量指标 $x \sim N(\mu,\sigma^2)$.以μ小者为佳, 厂方要求所确定的验收方案对高质量的产品$(\mu \le \mu_0)$能以高概率$1-\alpha$为买方所接受. 买方则要求低质产品$(\mu \ge \mu_0+\delta)$能以高概率$1-\beta$被拒绝. α,β由买方和厂方协商给出. 并采取一次抽样确定该批产品是否为买方所接受. 问应怎样安排抽验方案. 已知$\mu_0=120,\ \delta=20,\ \sigma^2=900,\ \alpha,\beta$均取$0.05$.

解　检验问题可表达为$H_0:\mu \le \mu_0, H_1:\mu > \mu_0,$



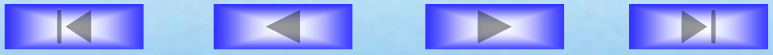
且要求当$\mu \geq \mu_1 + \delta$时能以$1-\beta = 0.95$的概率拒绝H_0.由 Z 检验,拒绝域为

$$\frac{\overline{x} - \mu_0}{\sigma / \sqrt{n}} \geq z_\alpha .$$

故OC函数为 $\beta(\mu) = P_\mu \left\{ \frac{\overline{X} - \mu_0}{\sigma / \sqrt{n}} < z_\alpha \right\}$

$$= P_\mu \left\{ \frac{\overline{X} - \mu}{\sigma / \sqrt{n}} < z_\alpha - \frac{\mu - \mu_0}{\sigma / \sqrt{n}} \right\}$$

$$= \Phi\left(z_\alpha - \frac{\mu - \mu_0}{\sigma / \sqrt{n}} \right). \quad (*)$$



现在要求当 $\mu \geq \mu_0 + \delta$ 时, $\beta(\mu) \leq \beta$.

因为 $\beta(\mu)$ 是 μ 的递减函数, 故只需 $\beta(\mu_0 + \delta) = \beta$.

此时, 由(*)式可得

$$\sqrt{n} \geq \frac{(z_\alpha + z_\beta)\sigma}{\delta}$$

根据给定的数据知 $n \geq 24.35$, 故取 $n = 25$.

即 $\bar{x} \geq 129.87$ 时, 买方就拒绝这批产品,

而当 $\bar{x} < 129.87$ 时, 买方就接受这批产品.

三、t 检验法的OC函数

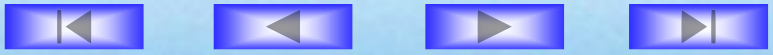
1. 右边检验问题

$H_0:\mu\le\mu_0, H_1:\mu>\mu_0$的$OC$函数是

$$\beta(\mu)=P_\mu(\text{接受}H_0)=P_\mu\left\{\frac{\overline{X}-\mu_0}{S/\sqrt{n}}<t_\alpha(n-1)\right\}$$

其中变量

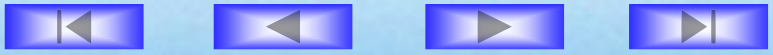
$$\frac{\overline{X}-\mu_0}{S/\sqrt{n}}=\left(\frac{\overline{X}-\mu}{\sigma/\sqrt{n}}+\lambda\right)\Big/\left(\frac{S}{\sigma}\right),\quad \lambda=\frac{\mu-\mu_0}{\sigma/\sqrt{n}}.$$



我们称变量 $\dfrac{\overline{X}-\mu_0}{S/\sqrt{n}}$ 服从非中心参数为 λ，自由度为$n-1$的非中心t分布.

当$\lambda=0$时, 它是通常的$t(n-1)$变量.

若给定α,β以及$\delta>0$, 则可从教材附表7查得所需容量n, 使得当$\mu\in H_1$且$\dfrac{\mu-\mu_0}{\sigma}\geq\delta$时, **犯第 II类错误的概率不超过** β.



2. 左边检验问题

假设检验 $H_0:\mu \geq \mu_0,\quad H_1:\mu < \mu_0.$

若给定 α,β 以及 $\delta > 0,$

则可从附表 6 查得所需容量 $n,$

使当 $\mu \in H_1$ 且 $\dfrac{\mu-\mu_0}{\sigma} \leq -\delta$ 时,

犯第 II类错误的概率不超过给定的 β.

3. 双边检验问题

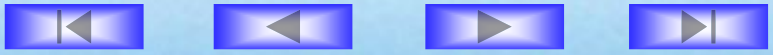
假设检验 $H_0:\mu=\mu_0,\quad H_1:\mu\neq\mu_0.$

若给定 α,β 以及 $\delta>0,$

则可从附表7查得所需容量 $n,$

使当 $\mu\in H_1$ 且 $\dfrac{|\mu-\mu_0|}{\sigma}\geq\delta$ 时,

犯第II类错误的概率不超过给定的 β.

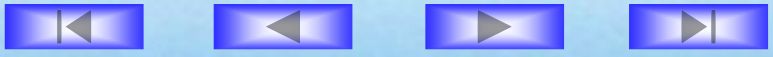


例2 考虑在显著水平$\alpha = 0.05$下进行 t 检验,

$$H_0 : \mu \le 68, \quad H_1 : \mu > 68,$$

(1) 要求在 H_1 中$\mu \ge \mu_1 = 68 + \sigma$时, 犯第II类错误的概率不超过 $\beta = 0.05$, 求所需样本容量.

(2) 若样本容量$n = 30$,问在 H_1 中$\mu = \mu_1 = 68 + 0.75\sigma$时, 犯第II类错误的概率是多少?



解 (1) $\alpha=\beta=0.05,\ \mu_0=68,$

附表5-1

$$\delta=\frac{\mu_1-\mu_0}{\sigma}=\frac{(68+\sigma)-68}{\sigma}=1,$$ 查表 7 知 $n=13.$

(2) 现在 $\alpha=0.05,\ n=30,$

$$\delta=\frac{\mu_1-\mu_0}{\sigma}=\frac{(68+0.75\sigma)-68}{\sigma}=0.75,$$

查表 7 知 $\beta=0.01$.

附表5 -2

例3 考虑在显著水平$\alpha = 0.05$下进行t检验,

$$H_0 : \mu = 14, \quad H_1 : \mu \neq 14,$$

要求在H_1中$\dfrac{|\mu - 14|}{\sigma} \geq 0.4$时,犯第II类错误的概率不超过$\beta = 0.1$, 求所需样本容量.

解 此处$\alpha = 0.05$, $\beta = 0.1$, $\delta = 0.4$,

查表 7 知$n = 68$.

附表5 -3

4. 求样本容量的一种近似方法

若只给出α,β及$|\mu_1-\mu_0|$,**怎样确定所需样本容量?**

先适当取一值n_1, 抽取容量为n_1的样本,

根据这一样本计算s^2的值,

以s^2作为σ^2的估计, 算出δ的近似值.

由α,β,δ的值查附表7定出样本的容量, 记为n_2 .

若$n_1 \geq n_2$, 则取n_1作为所求的容量, 即$n=n_1$.

否则, 再抽 $n_2 - n_1$ 个独立观察值与原来抽 得的观察值合并, 重新计算δ 的近似值.

然后用δ 的新近似值和α, β 查附表7, 再次定出样本容量, 记为n_3.

若 $n_2 \geq n_3$, 则取 n_2 作为所求的容量, 即$n = n_2$. **否则再按上述方法重复进行.**

一般, 只需试少数几次就可以得到所求的样本容量n.

5. 两个正态总体均值差的 t 检验问题

若两个正态总体$N(\mu_1,\sigma_1^2)$, $N(\mu_2,\sigma_2^2)$中

$$\sigma_1^2=\sigma_2^2=\sigma^2,\ (\sigma^2\text{未知})$$

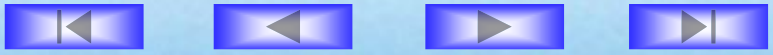
均值差$\mu_1-\mu_2$ 的检验问题

$$H_0:\mu_1-\mu_2=0,\ H_1:\mu_1-\mu_2\neq 0\ (>0\text{或}<0),$$

当分别自两个总体取得的相互独立的样本容量

$n_1=n_2=n$时，给定α,β 及 $\dfrac{|\mu_1-\mu_2|}{\sigma}$ 的值后，

可以查附表8得到所需的样本容量.



若两个正态总体 $N(\mu_1,\sigma_1^2)$, $N(\mu_2,\sigma_2^2)$中

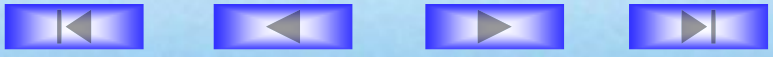
$$\sigma_1^2=\sigma_2^2=\sigma^2, \quad (\sigma^2\text{未知})$$

在均值差$\mu_1-\mu_2$的检验问题

$H_0:\mu_1-\mu_2=0$, $H_1:\mu_1-\mu_2\neq 0$ $(>0\text{或}<0)$,

的t检验法中,**当分别自两个总体取得的相互独立的样本容量**$n_1=n_2=n$**时,可以查附表8得到所需的样本容量.** 当$\mu\in H_1$且$(\mu-\mu_2)/\sigma\geq\delta$时, **犯第 II类错误的概率小于或等于** β.

当仅给出 α,β 以及$|\mu-\mu_2|$的值时, 可按类似于上面所说的方法处理.



例4 需比较两种汽车用的燃料的辛烷值, 得数据

燃料A	81	84	79	76	82	83	84	80	79	82	81	79
燃料B	76	74	78	79	80	79	82	76	81	79	82	78

燃料的辛烷值越高, 燃料质量越好. 因燃料 B 较燃料 A 价格便宜, 因此, 如果两者辛烷值相同时, 则使用燃料 B. 但若含量的均值差 $\mu_A - \mu_B \geq 5$, 则使用燃料 A. 设两总体的分布均可认为是正态的, 而两个样本相互独立. 问应采用那种燃料?

$(\alpha = \beta = 0.01)$

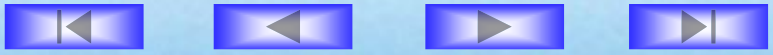
解 按题意需要在显著水平 $\alpha = 0.01$ 下检验假设

$$H_0:\mu_A-\mu_B\le 0,\quad H_1:\mu_A-\mu_B>0.$$

并要求 $\mu_A-\mu_B\ge 5$ 时, **犯第II类错误的概率不超过** $\beta=0.01$, **所取的样本容量** $n_A=n_B=12$, 且有 $\bar{x}_A=80.83$, $\sigma_A^2=\sigma_B^2$, $s_A^2=5.61$, $s_B^2=6.06$, **经水平为 0.1 的** F **检验知:** $\bar{x}_B=78.67$, 记为σ^2. 因 $n_1=n_2$, 取 $\hat{\sigma}^2=\dfrac{s_A^2+s_B^2}{2}=5.835$ 作为σ^2 的点估计, 取$\sigma=\sqrt{\hat{\sigma}^2}$, 于是 $\delta=\dfrac{5}{\hat{\sigma}}=2.07$, 查表,

当$\alpha=\beta=0.01$, $\delta=2.07$ 时 $n\ge 8$.

附表5-4



现 $n=12$, 故已近似地满足要求.

右边检验的拒绝域为

$$t=\frac{\bar{x}_A-\bar{x}_B}{s_w\sqrt{\frac{1}{n_A}+\frac{1}{n_B}}}\geq t_{0.01}(n_1+n_2-2)=2.5083\,.$$

由样本观察值算得 $t=2.19<2.5083$,

故接受 H_0, 即采用 B 种燃料.

补充例题

四、小结

两种检验法的OC函数如表

	右边检验	左边检验	双边检验
Z 检验	$\beta(\mu)=\Phi(z_\alpha-\lambda)$ $\lambda=\dfrac{\mu-\mu_0}{\sigma/\sqrt{n}}.$	$\beta(\mu)=\Phi(z_\alpha+\lambda)$ $\lambda=\dfrac{\mu-\mu_0}{\sigma/\sqrt{n}}.$	$\beta(\mu)=$ $\Phi(z_{\alpha/2}-\lambda)+\Phi(z_{\alpha/2}+\lambda)-1$ $\lambda=\dfrac{\mu-\mu_0}{\sigma/\sqrt{n}}.$
t 检验	$\beta(\mu)=$ $P_\mu\left\{\dfrac{\overline{X}-\mu_0}{S/\sqrt{n}}<t_\alpha(n-1)\right\}$ $\dfrac{\overline{X}-\mu_0}{S/\sqrt{n}}=$ $\left(\dfrac{\overline{X}-\mu}{\sigma/\sqrt{n}}+\lambda\right)\Big/\left(\dfrac{S}{\sigma}\right)$	$\beta(\mu)=$ $P_\mu\left\{\dfrac{\overline{X}-\mu_0}{S/\sqrt{n}}>-t_\alpha(n-1)\right\}$ $\dfrac{\overline{X}-\mu_0}{S/\sqrt{n}}=$ $\left(\dfrac{\overline{X}-\mu}{\sigma/\sqrt{n}}+\lambda\right)\Big/\left(\dfrac{S}{\sigma}\right)$	$\beta(\mu)=$ $P_\mu\left\{-t_{\alpha/2}(n-1)<\dfrac{\overline{X}-\mu_0}{S/\sqrt{n}}<t_{\alpha/2}(n-1)\right\}$ $\dfrac{\overline{X}-\mu_0}{S/\sqrt{n}}=$ $\left(\dfrac{\overline{X}-\mu}{\sigma/\sqrt{n}}+\lambda\right)\Big/\left(\dfrac{S}{\sigma}\right)$

第六节 分布拟合检验

一、χ^2 拟合检验法

二、分布族的χ^2拟合检验

三、偏度、峰度检验

四、小结

一、单个分布的 χ^2 拟合检验法

1. χ^2 检验法的定义

设总体X的分布未知，$x_1, x_2, \cdots, x_n$ 是来自X的样本值．我们来检验假设

H_0：总体 X 的分布函数为$F(x)$,

H_1：总体 X 的分布函数不是 $F(x)$,

其中设$F(x)$不含未知参数 .

说明

(1)在这里备择假设H_1可以不必写出.

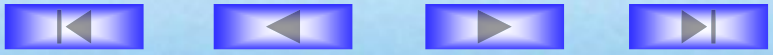
(2)若总体 X 为离散型：则上述假设相当于

H_0：总体 X 的分布律为$P\{X=t_i\}=p_i, i=1,2,\cdots.$

(3)若总体 X 为连续型： 则上述假设相当于

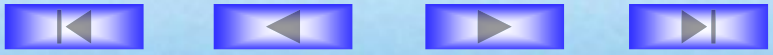
H_0：总体 X 的概率密度为 $f(x)$.

(4) 在使用χ^2检验法检验假设H_0时, 若$F(x)$的形式已知, 但其参数值未知，需要先用最大似然估计法估计参数, 然后作检验.



2. χ^2检验法的基本思想

将在H_0下X可能取值的全体Ω分成互不相交的子集$A_1, A_2, \cdots, A_k$，以$f_i (i = 1,2,\cdots,k)$记样本观察值$x_1, x_2, \cdots, x_n$中落在A_i的个数，这表示事件A_i = {X的值落在子集A_i内}在n次独立试验中发生f_i次，于是在这n次试验中事件A_i发生的频率为f_i/n. **另一方面**，当H_0为真，我们可以根据H_0中所假设的X的分布函数来计算事件A_i的概率，**得到** $p_i = P(A_i)$，$i = 1,2,\cdots,k$. 频率f_i/n与概率p_i会



有差异,但一般来说,若 H_0 为真,且试验次数又多时,**这种差异不应太大**,因此 $(\frac{f_i}{n}-p)^2$ 不应太大,

我们采用形如

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

的统计量来度量样本与 H_0 中所假设的分布的吻合程度,其中 C_i 为给定的常数 .

3.皮尔逊定理

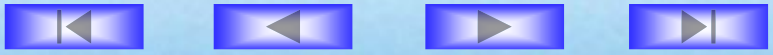
设检验假设H_0的统计量为

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

定理

若n充分大(≥ 50)，则H_0为真时统计量近似地服从$\chi^2(k-1)$分布.

据以上的讨论，当H_0为真时，式中的χ^2不应太大，如χ^2过分大就拒绝H_0，**拒绝域的形式为**



$$\chi^2 \geq G \qquad (G\text{为正常数}).$$

对于给定的显著性水平α, 确定G使

$$P\{\text{当}H_0\text{为真时拒绝}\ H_0\} = P_{H_0}\{\chi^2 \geq G\} = \alpha$$

由上述定理得$G = \chi^2_\alpha(k-1)$.

即当样本观察值使 式中的χ^2的值有

$$\chi^2 \geq \chi^2_\alpha(k-1),$$

则在显著性水平 α下拒绝 H_0; 否则就接受 H_0.

这就是单个分布的 χ^2拟和检验法 .

例1 下表列出了某一地区在夏季的一个月中由100个气象站报告的雷暴雨的次数.

i	0	1	2	3	4	5	≥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 的气象战数.**试用** χ^2拟合检验法检验雷暴雨的次数 X 是否服从均值 $\lambda = 1$的泊松分布(取显著性水平$\alpha = 0.05$).

解　按题意需检验假设

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

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

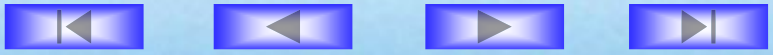
$$p_i=P\{X=i\}=\frac{e^{-1}}{i!},\quad i=0,1,\cdots 5.$$

例如 $p_0=P\{X=0\}=e^{-1}=0.36788$

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

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

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	
$A_4: \{X = 4\}$	6	$e^{-1}/24$	1.533	54.92
$A_5: \{X = 5\}$	2	$e^{-1}/120$	0.307	
$A_6: \{X = 6\}$	0	$1 - \sum_{i=0}^{5} p_i$	0.059	
			8.03	



计算结果如表所示, 其中有些$np_i < 5$的组予以适当合并,使得每组均有$np_i \geq 5$,**如表中第4列花括号所示,** 并组后$k = 4$,

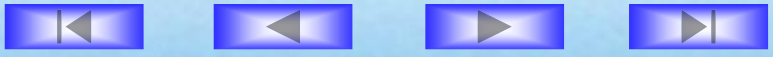
χ^2的自由度为 $k-1 = 4-1 = 3$

$\chi^2_{0.05}(k-1) = \chi^2_{0.05}(3) = 7.815$

现在

$$\chi^2 = 127.04 - 100 = 27.04 > 7.815$$

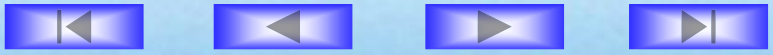
故在显著性水平0.05下拒绝H_0,认为样本不是来自均值 $\lambda = 1$的泊松分布.



例2　在研究牛的毛色与牛角的有无，这样两对性状分离现象时，用黑色无角牛与红色有角牛杂交，子二代出现黑色无角牛192头，黑色有角牛78头，红色无角牛72头，红色有角牛18头，共360头，问这两对性状是否符合孟德尔遗传规律中9：3：3：1的遗传比例?

解　现将题中的数据列表如下

序号	1	2	3	4
种类	黑色无脚	黑色有角	红色无脚	红色有角
数量	192	78	72	18
A_i	A_1	A_2	A_3	A_4



以X记各种牛的序号，按题意需检验各类牛的头数符合比例9：3：3：1，即(9/16):(3/16):(3/16):(1/16).

需检验假设：H_0：X的分布律为

X	1	2	3	4
p_k	9/16	3/16	3/16	1/16

取显著性水平为0.1. 所需计算列表如下(n=360)

A_i	f_i	p_i	np_i	$f_i^2/(np_i)$
A_1	192	9/16	$360\times 9/16 = 202.5$	$192^2/202.5 = 182.04$
A_2	78	3/16	$360\times 3/16 = 67.5$	$78^2/67.5 = 90.13$
A_3	72	3/16	$360\times 3/16 = 67.5$	$72^2/67.5 = 76.8$
A_4	18	1/16	$360\times 1/16 = 22.5$	$18^2/22.5 = 14.4$

现在

$$\chi^2 = 363.37 - 360 = 3.37,$$

$$k = 4, \quad \chi^2_{0.1}(4-1) = 6.251 > 3.37,$$

故接受H_0,认为两对性状是否符合孟德尔遗传规律中9：3：3：1的遗传比例.

二、 分布族的χ^2拟合检验

需检验的原假设是:

H_0 : 总体X的分布函数是$F(x;\theta_1,\cdots,\theta_r)$,

其中F的形式已知,而$\theta=(\theta_1,\theta_2,\cdots,\theta_r)$是未知参数,

它们在某一个范围取值. 在$F(x;\theta_1,\cdots,\theta_r)$中当参数$\theta_1,\theta_2,\cdots,\theta_r$取不同的值时 ,**就得到不同的分布.**

当H_0为真,由H_0所假设的分布函数来计算 $P(A_i)$,

得到 $P(A_i)=p_i(\theta_1,\theta_2,\cdots,\theta_r)=p_i(\theta)=p_i$.

求出p_i的估计值 $\hat{p}_i = \hat{P}(A_i)$,

$$\chi^2 = \sum_{i=1}^{k} \frac{f_i^2}{n\hat{p}_i} - n$$

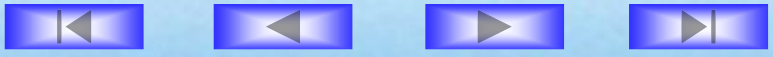
作为检验假设 H_0的统计量. **可以证明, 在某些条件下,** 在H_0为真时近似地有

$$\chi^2 = \sum_{i=1}^{k} \frac{f_i^2}{n\hat{p}_i} - n \sim \chi^2(k-r-1)$$

与在一中一样可得假设检验问题的拒绝域为

$$\chi^2 \geq \chi_\alpha^2(k-r-1),$$

α为显著性水平.

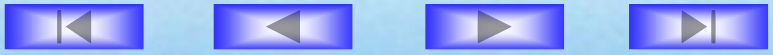


例3　在一实验中，每隔一定时间观察一次由某种铀所放射的到达计数器上的α粒子数X，共观察了100次，得结果如下表所示.

i	0	1	2	3	4	5	6	7	8	9	10	11	≥12
f_i	1	5	16	17	26	11	9	9	2	1	2	1	0
A_i	A_0	A_1	A_2	A_3	A_4	A_5	A_6	A_7	A_8	A_9	A_{10}	A_{11}	A_{12}

其中f_i是观察到有i个α粒子的次数 **.从理论上考虑知X应服从泊松分布**

$$P\{X=i\}=\frac{\lambda^i e^{-\lambda}}{i!},\qquad i=0,1,2,\cdots.$$



问上式是否符合实际 (取$\alpha = 0.05$)? **即在显著性水平0.05下检验假设**

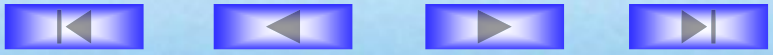
H_0 : 总体X服从泊松分布,

$$P\{X = i\} = \frac{\lambda^i e^{-\lambda}}{i!}, \qquad i = 0,1,2,\cdots.$$

解 因在H_0中参数λ未具体给出，所以先估计λ.

由最大似然估计法得: $\hat{\lambda} = \bar{x} = 4.2$.在H_0假设下,

即在X服从泊松分布的假设下，X所有可能取得值为 $\Omega = \{0,1,2,\cdots\}$, 将$\Omega$分成如表所示的两两不相交的子集$A_0, A_1, \cdots, A_{12}$. 则$P\{X = i\}$有估计



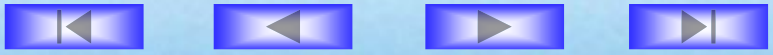
$$\hat{p}_i = \hat{P}\{X = i\} = \frac{4.2^i e^{-4.2}}{i!}, \quad i = 0,1,2,\cdots.$$

例如

$$\hat{p}_0 = \hat{P}\{X = 0\} = e^{-4.2} = 0.015,$$

$$\hat{p}_3 = \hat{P}\{X = 3\} = \frac{4.2^3 e^{-4.2}}{3!} = 0.185,$$

$$\hat{p}_{12} = \hat{P}\{X \geq 12\} = 1 - \sum_{i=0}^{11} \hat{p}_i = 0.002.$$



例3的 χ^2 拟合检验计算表

A_i	f_i	$\hat{p}_i$	$n\hat{p}_i$	$f_i^2 / n\hat{p}_i$
A_0	1	0.015	1.5	4.615
A_1	5 (6)	0.063 (0.078)	6.3 (7.8)	
A_2	16	0.132	13.2	19.394
A_3	17	0.185	18.5	15.622
A_4	26	0.194	19.4	34.845
A_5	11	0.163	16.3	7.423
A_6	9	0.114	11.4	7.105
A_7	9	0.069	6.9	11.739
A_8	2	0.036	3.6	5.538
A_9	1	0.017	1.7	
A_{10}	2 (6)	0.007 (0.065)	0.7 (6.5)	
A_{11}	1	0.003	0.3	$\sum$ =106.281
A_{12}	0	0.002	0.2	

其中有些$n\hat{p}_i < 5$的组予以合并, 使得每组均有$n\hat{p}_i \geq 5$, 如表中第四列花括号所示.

并组后$k = 8$, 故 χ^2 的自由度为$8-1-1=6$,

$$\chi_\alpha^2(k-r-1) = \chi_{0.05}^2(6) = 12.592 > 6.2815,$$

故接受 H_0, 认为样本来自泊松分布总体.

例4　自1965年1月1日至1971年2月9日共2231天中，全世界记录到里氏震级4级和4级以上地震计162次，统计如下：

相继两次地震间隔天数x	0~4	5~9	10~14	15~19	20~24	25~29	30~34	35~39	≥ 40
出现的频数	50	31	26	17	10	8	6	6	8

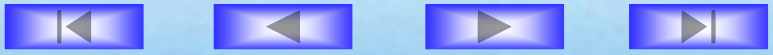
试检验相继两次地震间隔的天数 X 服从指数分布 $(\alpha = 0.05)$.

解　按题意检验假设

H_0：X 的概率密度为

$$f(x)=\begin{cases}\dfrac{1}{\theta}e^{-x/\theta}, & x>0,\\ 0, & x\le 0.\end{cases}$$

在这里， H_0 中的参数 θ 未给出，**先求最大似然**估计法求得 θ 的估计为 $\hat{\theta}=\bar{x}=\dfrac{2231}{162}=13.77$. 在 H_0 **下，** X 可能取值的全体 Ω 为区间 $[0,\infty)$. 将区间 $[0,\infty)$ 分为 $k=9$ 个互不重叠的小区间: $A_1=[0,4.5]$,



$A_2=(4.5,9.5],\cdots,A_9=(39.5,\infty)$，**如教材中表8.6第二列所示.** 若$H_0$为真，$X$的分布函数的估计为

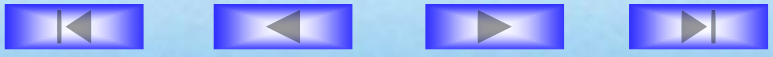
$$\hat{F}(x)=\begin{cases}1-e^{-x/13.77}, & x>0,\\ 0, & x\le 0.\end{cases}$$

由上式可得概率$p_i=P(A_i)$的估计：

$$\hat{p}_i=\hat{P}(A_i)=\hat{P}\{a_i<X\le a_{i+1}\}=\hat{F}(a_{i+1})-\hat{F}(a_i).$$

例如

$$\hat{p}_2=\hat{P}(A_2)=\hat{P}\{4.5<X\le 9.5\}$$



$$\hat{p}_2=\hat{P}(A_2)=\hat{P}\{4.5<X\le 9.5\}$$

$$=\hat{F}(9.5)-\hat{F}(4.5).$$

$$=0.2196,$$

而

$$\hat{P}_9=\hat{P}(A_2)=1-\sum_{i=1}^{8}\hat{F}(A_i)=0.0568.$$

将计算结果列表如下:

例4的 χ^2 拟合检验计算表

A_i	f_i	$\hat{p}_i$	$n\hat{p}_i$	$f_i^2 / n\hat{p}_i$
$A_1: 0 \le x \le 4.5$	50	0.2788	45.1656	55.3519
$A_2: 4.5 < x \le 9.5$	31	0.2196	35.5752	27.0132
$A_3: 9.5 < x \le 14.5$	26	0.1527	24.7374	27.3270
$A_4: 14.5 < x \le 19.5$	17	0.1062	17.2044	16.7980
$A_5: 19.5 < x \le 24.5$	10	0.0739	11.9718	8.3530
$A_6: 24.5 < x \le 29.5$	8	0.0514	8.3268	7.6860
$A_7: 29.5 \le x \le 34.5$	6	0.0358	5.7996	6.2073
$A_8: 34.5 < x \le 39.5$	6	0.0248	4.0176	14.8269
$A_9: 39.5 < x < \infty$	8	0.0568	9.2016	13.2192
				$\sum = 163.5633$

现在$\chi^2 = 163.5633 - 162 = 1.5633$, 因为

$$\chi^2_{0.05}(k-r-1)=\chi^2_{0.05}(8-1-1)=\chi^2_{0.05}(6)=12.592$$

$$> 1.5633$$

故在显著性水平0.05下接受H_0, 认为X 服从指数分布.

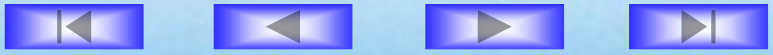
例5　对于第六章§2例1中的数据，　试检验它们是否来自正态总体X (取显著水平$\alpha = 0.1$).

解　需检验假设

H_0 : X 的概率密度为

$$f(x) = \frac{1}{\sqrt{2\pi}\sigma} e^{-\frac{(x-\mu)^2}{2\sigma^2}}, \quad -\infty < x < \infty,$$

因在H_0中未给出μ, σ^2的数值. 需先估计μ, σ^2. **由**最大似然估计法得μ, σ^2的估计值为$\hat{\mu} = 143.8$, $\hat{\sigma}^2 = (6.0)^2$. 我们将在H_0下X可能取值的区间



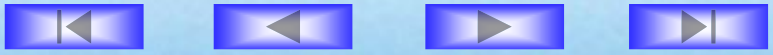
$(-\infty,\infty)$分为7个小区间, 并取事件A_i, **如教材中表8.7第一列所示.**若H_0为真X的概率密度的估计为

$$\hat{f}(x)=\frac{1}{\sqrt{2\pi}\times 6}e^{-\frac{(x-143.8)^2}{2\times 6^2}},\qquad -\infty<x<\infty,$$

按上式并查标准正态分布的分布函数表即可得概率$P(A_i)$的估计. **例如**

$$\hat{p}_2=\hat{P}(A_2)=\hat{P}\{129.5<X\le 134.5\}$$

$$=\Phi\left(\frac{134.5-143.8}{6}\right)-\Phi\left(\frac{129.5-143.8}{6}\right)$$



$$\hat{p}_2=\hat{P}(A_2)=\hat{P}\{129.5<X\le 134.5\}$$

$$=\Phi\left(\frac{134.5-143.8}{6}\right)-\Phi\left(\frac{129.5-143.8}{6}\right)$$

$$=\Phi(-1.55)-\Phi(-2.38)=0.0519.$$

将计算结果列表如下:

例5的 χ^2 拟合检验计算表

A_i	f_i	$\hat{p}_i$	$n\hat{p}_i$	$f_i^2 / n\hat{p}_i$
$A_1 : x \le 129.5$	1	0.0087	0.73	4.91
$A_2 : 129.5 < x \le 134.5$	4 (5)	0.0519	4.36 (5.09)	
$A_3 : 134.5 < x \le 139.5$	10	0.1752	14.72	6.79
$A_4 : 139.5 < x \le 144.5$	33	0.3120	26.21	41.55
$A_5 : 144.5 < x \le 149.5$	24	0.2811	23.61	24.4
$A_6 : 149.5 < x \le 154.5$	3	0.1336	11.22	10.02
$A_7 : 154.5 \le x \le \infty$	9 (12)	0.0375	3.15 (14.37)	

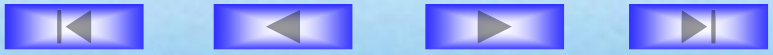
现在$\chi^2 = 87.67 - 84 = 3.67$，因为

$\chi^2_{0.1}(k-r-1)=\chi^2_{0.1}(5-2-1)=\chi^2_{0.1}(2)=4.605>3.67$，

故在水平0.1下接受H_0，即认为数据来自正态分布总体.

三、偏度、峰度检验

正态分布随机变量较广泛地存在于客观世界，因此,当研究一连续型总体时,人们往往先考察它是是否服从正态分布.上面介绍的χ^2拟合检验法虽然是检验总体分布的较一般的方法,但用它来检验总体的正态性时,犯第II类错误的概率往往较大.为此在对检验正态总体的种种方法进行比较后，认为“偏度、峰度检验法”和“夏皮罗 - 威尔克法”较为有效.(此处只介绍前一种)



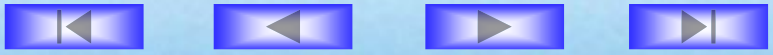
1. 随机变量的偏度和峰度的定义

随机变量 X 的偏度和峰度分别指 X 的标准化变量 $\dfrac{X-E(X)}{\sqrt{D(X)}}$ 的三阶中心矩和四阶中心矩：

$$\nu_1 = E\left[\left(\frac{X-E(X)}{\sqrt{D(X)}}\right)^3\right] = \frac{E[(X-E(X))^3]}{(D(X))^{3/2}},$$

$$\nu_2 = E\left[\left(\frac{X-E(X)}{\sqrt{D(X)}}\right)^4\right] = \frac{E[(X-E(X))^4]}{(D(X))^2}.$$

当随机变量 X 服从正态分布时，$\nu_1 = 0$ 且 $\nu_2 = 3$.



2. 样本偏度和样本峰度的定义

设 $X_1, X_2, \cdots, X_n$ 是来自总体 X 的样本,

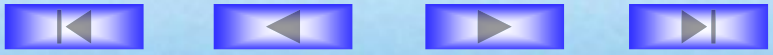
则 $G_1 = \dfrac{B_3}{B_2^{3/2}}, G_2 = \dfrac{B_4}{B_2^2}$ 分别称为样本偏度和样 本峰度.

其中 $B_k (k = 2,3,4)$ 是样本 k 阶中心矩.

若总体 X 为正态变量, 则当 n 充分大时, 近似地有

$$G_1 \sim N\left(0, \frac{6(n-2)}{(n+1)(n+3)}\right),$$

$$G_2 \sim N\left(3 - \frac{6}{n+1}, \frac{24n(n-2)(n-3)}{(n+1)^2(n+3)(n+5)}\right).$$



3. 偏度、峰度检验法

设 $X_1, X_2, \cdots, X_n$ 是来自总体 X 的样本,

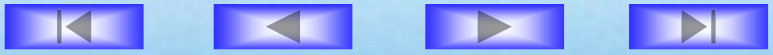
现在来检验假设 H_0：X 为正态总体.

记 $\sigma_1 = \sqrt{\dfrac{6(n-2)}{(n+1)(n+3)}}, \sigma_2 = \sqrt{\dfrac{24n(n-2)(n-3)}{(n+1)^2(n+3)(n+5)}},$

$$\mu_2 = 3 - \frac{6}{n+1}, \quad U_1 = \frac{G_1}{\sigma_1}, \quad U_2 = \frac{G_2 - \mu_2}{\sigma_2}.$$

当 H_0 为真且 n 充分大时, 近似地有

$$U_1 \sim N(0,1), \ U_2 \sim N(0,1).$$

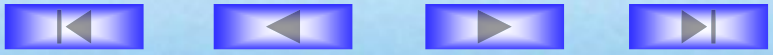


由第六章第三节知样本偏度 G_1 和样本峰度 G_2 分别依概率收敛于总体偏度 v_1 和总体峰度 v_2. **因此**当 H_0 为真且 n 充分大时, 一般地 G_1 与 $v_1=0$ 的偏离**不应太大**, G_2 与 $v_2=3$ 的偏离不应太大. **故从直观来**看当 $|u_1|$ 或 $|u_2|$ 过大时, 就拒绝 H_0. **取显著性水平**为 α, H_0 的拒绝域为

$$|u_1|\geq k_1 \quad \text{或} \quad |u_2|\geq k_2, \qquad (*)$$

其中 k_1 和 k_2 由下两式决定,

$$P_{H_0}\{|U_1|\geq k_1\}=\frac{\alpha}{2}; \qquad P_{H_0}\{|U_2|\geq k_2\}=\frac{\alpha}{2}.$$



这里记号$P_{H_0}\{\cdot\}$表示当H_0为真时事件$\{\cdot\}$的概率，

即$k_1 = z_{\alpha/4}$，$k_2 = z_{\alpha/4}$，**于是得拒绝域为**

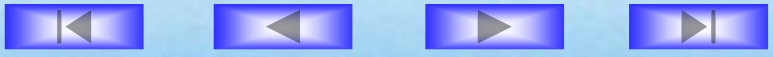
$$|u_1| \geq z_{\alpha/4} \quad \text{或} \quad |u_2| \geq z_{\alpha/4}.$$

下面验证当n充分大时上述检验法近似地满足显著水平为α的要求. 事实上当n充分大时有

$$P\{\text{当}H_0\text{为真拒绝}H_0\}$$

$$= P_{H_0}\{(|U_1| \geq z_{\alpha/4}) \cup (|U_2| \geq z_{\alpha/4})\}$$

$$\leq P_{H_0}\{(|U_1| \geq z_{\alpha/4})\} + P_{H_0}\{(|U_2| \geq z_{\alpha/4})\}$$



$$P\{\text{当}H_0\text{为真拒绝}H_0\}$$

$$=P_{H_0}\{(|U_1|\geq z_{\alpha/4})\cup(|U_2|\geq z_{\alpha/4})\}$$

$$\leq P_{H_0}\{(|U_1|\geq z_{\alpha/4})\}+P_{H_0}\{(|U_2|\geq z_{\alpha/4})\}$$

$$=\frac{\alpha}{2}+\frac{\alpha}{2}=\alpha$$

例6 试用偏度、峰度检验法检验本节例5中的数据是否来自正态总体? $(\text{取}\alpha = 0.1)$

解 现在来检验假设

$$H_0: \text{数据来自正态总体.}$$

这里 $\alpha = 0.1,\ n = 84,$

$$\sigma_1 = \sqrt{\frac{6(n-2)}{(n+1)(n+3)}} = 0.2579$$

$$\sigma_2 = \sqrt{\frac{24n(n-2)(n-3)}{(n+1)^2(n+3)(n+5)}} = 0.4892,$$

$$\mu_2 = 3 - \frac{6}{n+1} = 2.9294,$$

下面来计算样本中心矩$B_k\,(k = 2,3,4)$

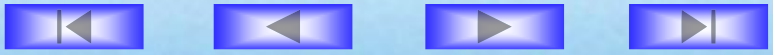
$$B_2 = A_2 - A_1^2,$$

$$B_3 = A_3 - 3A_2A_1 + 2A_1^3,$$

$$B_4 = A_4 - 4A_3A_1 + 6A_2A_1^{\,2} - 3A_1^4,$$

其中 $A_k = \frac{1}{n}\sum_{i=1}^{n} x_i^{\,k}\,(k = 1,2,3,4)$ 为k阶样本矩,计算得

$$A_1 = 143.7738,\quad A_2 = 20706.13,\quad A_3 = 2987099,$$



$A_4 = 4.316426\times10^8,\quad B_2 = 35.2246,\quad B_3 = -28.5,$

$B_4 = 3840,$

样本偏度和样本峰度的观察值为

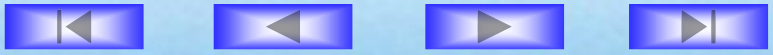
$$g_1 = -0.1363,\qquad g_2 = 3.0948,$$

而 $z_{\alpha/4} = z_{0.025} = 1.96$. **由(*)式, 拒绝域为**

$$|u_1| \geq |g_1/\sigma_1| \geq 1.96 \quad \text{或} \quad |u_2| = |g_2 - \mu_2| \geq 1.96.$$

现算得 $|u_1| = 0.5285 < 1.96,\ |u_2| = 0.3381 < 1.96,$ **故接受** H_0, **认为数据来自正态分布的总体.**

补充例题



四、小结

χ^2 检验法的定义、基本思想.

皮尔逊定理

偏度、峰度的定义及检验法.

第七节 秩和检验

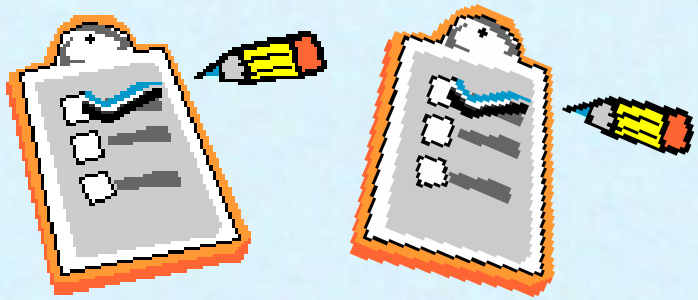
一、基本概念

二、典型例题

三、小结

一、基本概念

1.假设中的等价问题



设有两个连续型总体，它们的概率密度函数分别为 $f_1(x)$, $f_2(x)$ (均为未知)，

但已知 $f_1(x) = f_2(x-a)$, a为未知常数. (7.1)

要检验的各假设为

$$H_0: a = 0, \quad H_1: a < 0. \quad (7.2)$$

$$H_0: a = 0, \quad H_1: a > 0. \quad (7.3)$$

$$H_0: a = 0, \quad H_1: a \neq 0. \quad (7.4)$$

特别,若两个总体的均值存在,分别记为μ_1,μ_2,由于 f_1, f_2最多只差一平移, 则有 $\mu_2 = \mu_1 - a.$

此时, 上述各假设分别等价于

$$H_0: \mu_1 = \mu_2, \quad H_1: \mu_1 < \mu_2. \quad (7.2)'$$

$$H_0: \mu_1 = \mu_2, \quad H_1: \mu_1 > \mu_2. \quad (7.3)'$$

$$H_0: \mu_1 = \mu_2, \quad H_1: \mu_1 \neq \mu_2. \quad (7.4)'$$

2.秩的定义

秩 X 为一总体, 将容量为n的样本观察值按自小到大的次序编号排列成

$$x_{(1)} < x_{(2)} < \cdots < x_{(n)},$$

称 $x_{(i)}$ 的足标 i 为 $x_{(i)}$ 的秩, $i = 1,2,\cdots,n.$

例如:某旅行团人员的行李重量数据如表

重量(kg)	34	39	41	28	33

写出重量33的秩.

因为 $28 < 33 < 34 < 39 < 41$, **故33的秩为2.**

特殊情况:

如果在排列大小时出现了相同大小的观察值,则其秩的定义为足标的平均值.

例如:抽得的样本观察值按次序排成0,1,1,1,2,3,3,

则3个1的秩均为$\frac{2+3+4}{3}=3$,

两个3的秩均为$\frac{6+7}{2}=6.5$.

3.秩和的定义

现设1,2两总体分别抽取容量为n_1, n_2的样本,且设两样本独立. 这里总假定$n_1 \le n_2$.我们将这$n_1 + n_2$个观察值放在一起,按从小到大的次序排列,求出每个观察值的秩,然后将属于第1个总体的样本观察值的秩相加, 其和记为R_1,称为第1样本的秩和,其余观察值的的总和记作R_2,称为第2样本的秩和.显然, R_1和 R_2是离散型随机变量, 且有

$$R_1 + R_2 = \frac{1}{2}(n_1 + n_2)(n_1 + n_2 + 1).$$

4.秩和检验法的定义

秩和检验法是一种非参数检验法，它是一种用样本秩来代替样本值的检验法.

用秩和检验法可以检验两个总体的分布函数是否相等的问题.

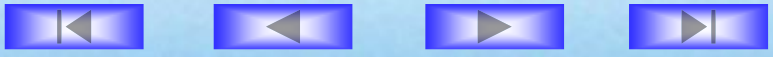
5. 双边检验 $H_0: a = 0, H_1: a \neq 0$ 问题

分析: 当H_0为真时, $f_1(x) = f_2(x)$,

此时, 两个独立样本实际上来自同一个总体.

因而第1样本中诸元素的秩应该随机地、分散地在自然数$1 \sim n_1 + n_2$中取值, 一般来说不应该过分集中取较小的或较大的值. 考虑到

$$\frac{1}{2}n_1(n_1+1) \leq R_1 \leq \frac{1}{2}n_1(n_1+2n_2+1),$$



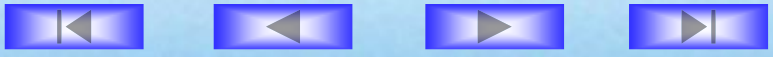
即知当H_0为真时秩和R_1一般来说不应该取太靠近上述不等式两端的值.

因而,当R_1的值r_1过分大或过分小时，我们都拒绝H_0. **据以上分析，对于双边检验问题(7.4).**

在给定显著性水平α下，H_0的拒绝域为

$$R_1 \le C_U\left(\frac{\alpha}{2}\right) \text{或} R_1 \ge C_L\left(\frac{\alpha}{2}\right),$$

其中临界点$C_U\left(\frac{\alpha}{2}\right)$是满足$P_{a=0}\left\{R_1 \le C_U\left(\frac{\alpha}{2}\right)\right\} \le \frac{\alpha}{2}$



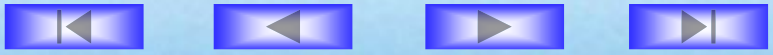
的最大整数, 而

$$C_L\left(\frac{\alpha}{2}\right)$$ 是满足 $$P_{a=0}\left\{R_1 \geq C_L\left(\frac{\alpha}{2}\right)\right\} \leq \frac{\alpha}{2}$$

的最小整数, 而犯第I类错误的概率是

$$P_{a=0}\left\{R_1 \leq C_U\left(\frac{\alpha}{2}\right)\right\} + P_{a=0}\left\{R_1 \geq C_L\left(\frac{\alpha}{2}\right)\right\}$$

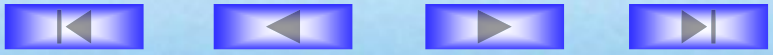
$$\leq \frac{\alpha}{2} + \frac{\alpha}{2} = \alpha.$$



6.求临界点的方法

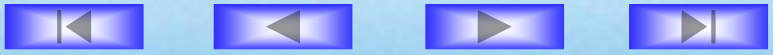
如果知道 R_1 的分布,则$C_U\left(\frac{\alpha}{2}\right), C_L\left(\frac{\alpha}{2}\right)$不难求得.下面以$n_1=3,\ \ n_2=4$的情况为例来说明求临界点的方法.

当$n_1=3,\ \ n_2=4$时, 第1个样本中各观察值的秩的不同取法共有$\binom{3+4}{3}=35$种, **现将这35种情况列表如下:**



秩	R_1	秩	R_1	秩	R_1	秩	R_1	秩	R_1
123	6	136	10	167	14	247	13	356	14
124	7	137	11	234	9	256	13	357	15
125	8	145	10	235	10	257	14	367	16
126	9	146	11	236	11	267	15	456	15
127	10	147	12	237	12	345	12	457	16
134	8	156	12	245	11	346	13	467	17
135	9	157	13	246	12	347	14	567	18

由于这35种情况的出现是等可能的, 由上表可求得 R_1 的分布律和分布函数如下表:



R_1的分布律和分布函数

R_1	6	7	8	9	10	11	12
$P\{R_1 = r_1\}$	1/35	1/35	2/35	3/35	4/35	4/35	5/35
$P\{R_1 \le r_1\}$	1/35	2/35	4/35	7/35	11/35	15/35	20/35

R_1	13	14	15	16	17	18	
$P\{R_1 = r_1\}$	4/35	4/35	3/35	2/35	1/35	1/35	
$P\{R_1 \le r_1\}$	24/35	28/35	31/35	33/35	34/35	1	

对于不同的α值,**容易写出检验问题的临界点和拒绝域.**

例如,给定$\alpha=0.2$, **由上表可知** 上表

$$P_{a=0}\{R_1 \le 7\} = \frac{2}{35} = 0.057 < 0.1 = \frac{\alpha}{2},$$

$$P_{a=0}\{R_1 \ge 17\} = \frac{2}{35} = 0.057 < 0.1 = \frac{\alpha}{2},$$

即有$C_U(0.1)=7, C_L(0.1)=17$. 故当$n_1=3$, $n_2=4$时,

在显著性水平0.2下检验问题的拒绝域为:

$$r_1 \le 7 \text{ 或 } r_1 \ge 17.$$

犯第I类错误的概率是

$$P_{a=0}\left\{R_1 \le C_U\left(\frac{\alpha}{2}\right)\right\} + P_{a=0}\left\{R_1 \ge C_L\left(\frac{\alpha}{2}\right)\right\}$$

$$= P_{a=0}\{R_1 \le 7\} + P_{a=0}\{R_1 \ge 17\}$$

$$= 0.057 + 0.057$$

$$= 0.114.$$

7.左边检验和右边检验

左边检测(7.2)的拒绝域为

(显著性水平为α)

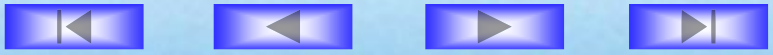
$$r_1 \le C_U(\alpha),$$

此处,临界点$C_U(\alpha)$是满足$P_{a=0}\{R_1 \le C_U(\alpha)\} \le \alpha$的最大整数.

而右边检验问题的拒绝域（显著性水平为α）

$$r_1 \ge C_L(\alpha),$$

此处,临界点$C_L(\alpha)$是满足$P_{a=0}\{R_1 \ge C_L(\alpha)\} \le \alpha$的最小整数.



例如，若给定$\alpha = 0.1$，抽取的样本容量为$n_1 = 3$，$n_2 = 4$，

则由上表知检验问题的拒绝域为

$$r_1 \geq 17.$$

此时犯第I类错误的概率为$2/35 < 0.1$.

书末附表9中列出了n_1和n_2自2到10为止的n_1, n_2的各种组合的临界点，以及相应的犯第I类错误的概率.

8.特殊情况(1)

可以证明当H_0为真时(即$a=0$),

$$\mu_{R_1}=E(R_1)=\frac{1}{2}n_1(n_1+n_2+1),$$

$$\sigma_{R_1}^{\ 2}=D(R_1)=\frac{1}{12}n_1n_2(n_1+n_2+1).$$

而当$n_1,n_2\geq 10$, H_0为真时, 近似地

$$R_1\sim N(\mu_{R_1},\sigma_{R_1}^2).$$

因此当$n_1, n_2 \geq 10$时,

选$Z = \dfrac{R_1 - \mu_{R_1}}{\sigma_{R_1}}$作检验统计量,

在显著性水平α下,

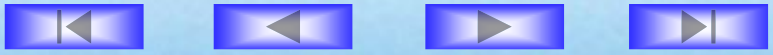
双边检验、右边检验、左边检验的拒绝域分别是

$$|Z| > z_{\alpha/2}, \quad Z > z_{\alpha}, \quad Z < -z_{\alpha}.$$

9.特殊情况(2)

将两个样本$n_1+n_2=n$个元素按自小到大的次序排列，若出现秩相同的组，设其中有t_i个数的秩为a_i，$i=1,2,\cdots,k$，$a_1<\cdots<a_k$，则当H_0为真时R_1的均值仍为

$$\mu_{R_1}=\frac{n_1(n_1+n_2+1)}{2},$$



而 R_1 的方差修正为

$$\sigma_{R_1}^{\ 2}=\frac{n_1n_2\left[n(n^2-1)-\sum_{i=1}^{k}t_i(t_i^2-1)\right]}{12n(n-1)}. \qquad (\#)$$

当 k 不大时，附表9仍能使用，但表中为近似值.

而当 $n_1, n_2 \geq 10$, H_0 为真，且 k 不大时，

近似地有 $R_1 \sim N(\mu_{R_1}, \sigma_{R_1}^2)$,

其中 $\mu_{R_1}=\dfrac{n_1(n_1+n_2+1)}{2}$，$\sigma_{R_1}{}^2$见上页(#)式.

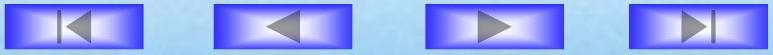
此时选 $\boxed{Z=\dfrac{R_1-\mu_{R_1}}{\sigma_{R_1}}}$ 作检验统计量，

来检验假设检验问题

$$H_0:\mu_1=\mu_2,\quad H_1:\mu_1<\mu_2.$$

$$H_0:\mu_1=\mu_2,\quad H_1:\mu_1>\mu_2.$$

$$H_0:\mu_1=\mu_2,\quad H_1:\mu_1\neq\mu_2.$$



二、典型例题

例1　为查明某种血清是否会抑制白血病, 选取患白血病已到晚期的老鼠9只, 其中有5至接受这种治疗, 另4 只则不作这种治疗. 设两样本相互独立.

从试验开始时计算, 其存活时间(以月计)如下:

不作治疗	1.9	0.5	0.9	2.1	
接受治疗	3.1	5.3	1.4	4.6	2.8

设治疗与否的存活时间的概率密度至多只差一个平移. 取 $\alpha = 0.05$, 问这种血清对白血病是否有抑制作用?

解　根据题意需检验老鼠的存活期是否有增长,

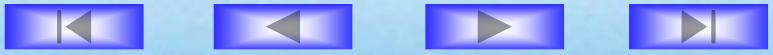
分别用μ_1和μ_2表示不作治疗和接受治疗的老鼠的存活时间总体的均值　,**需要检验的假设是:**

$$H_0:\mu_1=\mu_2,\ H_1:\mu_1<\mu_2.$$

这里 $n_1=4,\ n_2=5,\ \alpha=0.05$.

先求对应于$n_1=4$的一组观察值的秩和,**将两组数据放在一起按自小到大的次序排列,**

(对来自第1总体$(n_1=4)$的数据下面加_表示)



数据	0.5	0.9	1.4	1.9	2.1	2.8	3.1	4.6	5.3
秩	1	2	3	4	5	6	7	8	9

所以 R_1 的观察值为 $r_1 = 1+2+4+5 = 12.$

附表8

查附表8 知, $C_U(0.05) = 12$, 即拒绝域为 $r_1 \le 12$.

而现在 $r_1 = 12$, **故拒绝** H_0.

认为这种血清对白血病有抑制作用.

例2 某商店为了确定向公司 A 或公司 B 购买某种商品,将 A, B 公司以往各次进货的次品率进行比较,数据如下, 设两样本独立. 问两公司的商品的质量有无明显差异. 设两公司的商品的次品率的密度最多只差一个平移.

A	7.0	3.5	9.6	8.1	6.2	5.1	10.4	4.0	2.0	10.5			
B	5.7	3.2	4.2	11.0	9.7	6.9	3.6	4.8	5.6	8.4	10.1	5.5	12.3

解 分别用 μ_A 和 μ_B 记公司 A, B 的商品次品率总体的均值,

需要检验的假设是:

$$H_0:\mu_A=\mu_B,\quad H_1:\mu_A\neq\mu_B.$$

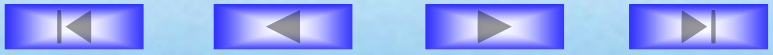
先将数据按大小次序排列, 得到对应于$n_1=10$样本的秩和为

$$r_1=1+3+5+8+12+14+15+17+20+21=116,$$

又当H_0为真时,

$$E(R_1)=\frac{1}{2}n_1(n_1+n_2+1)=\frac{1}{2}\times 10(10+13+1)=120,$$

$$D(R_1)=\frac{1}{12}n_1n_2(n_1+n_2+1)=260,$$



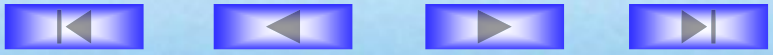
故当H_0为真时近似地有 $R_1 \sim N(120, 260)$

拒绝域为 $\dfrac{|R_1 - 120|}{\sqrt{260}} \geq z_{0.025} = 1.96,$

现在 R_1 的观察值为$r_1 = 116,$

$\dfrac{|r_1 - 120|}{\sqrt{260}} = 0.25 < 1.96$, 故接受H_0.

认为两个公司商品的质量无显著差异.



例3　两位化验员各自读得某种液体粘度如下:

化验员A	82	73	91	84	77	98	81	79	87	85
化验员B	80	76	92	86	74	96	83	79	80	75	79

设数据可以认为来自仅均值可能有差异的总体的样本. 在显著性水平$\alpha = 0.05$下检验假设

$$H_0 : \mu_1 = \mu_2, \quad H_1 : \mu_1 > \mu_2.$$

其中μ_1, μ_2分别为两总体的均值.

解　将两样本的元素混合,

按自小到大次序排列.

并求出各元素的秩如下.

数据	73	74	75	76	77	79	79	79	80	80	81
秩	1	2	3	4	5	7	7	7	9.5	9.5	11
数据	82	83	84	85	86	87	91	92	96	98	
秩	12	13	14	15	16	17	18	19	20	21	

现在 $n_1 = 10,\ n_2 = 11,\ n = 21,\ \alpha = 0.05$.

$$\mu_{R_1} = \frac{n_1(n_1+n_2+1)}{2} = \frac{10\times 22}{2} = 110,$$

$$k = 2,\quad \sum_{i=1}^{k} t_i(t_i^2-1) = 3\times(9-1)+2\times(4-1) = 30,$$

按教材中(7.9)式得 $\sigma_{R_1}^2 = 201.$

故当H_0为真时近似地有

$$R_1 \sim N(110, 201)$$

拒绝域为$\dfrac{|r_1 - 110|}{\sqrt{201}} \geq z_{0.05} = 1.645,$

现在R_1的观察值为$r_1 = 121,$

$\dfrac{|r_1 - 110|}{\sqrt{201}} = 0.776 < 1.645,$ 故接受H_0.

认为两个化验员所测得的数据无显著差异.

三、小结

基本概念:

秩的定义、秩和的定义、秩和检验法的定义.

左边检验 $H_0: a=0, \quad H_1: a<0$ 的拒绝域为

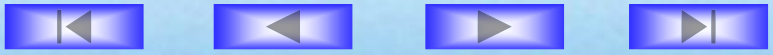
$r_1 \le C_U(\alpha)$, (显著性水平为 α)

右边检验 $H_0: a=0, \quad H_1: a>0$ 的拒绝域为

$r_1 \ge C_L(\alpha)$, (显著性水平为 α)

双边检验 $H_0: a=0, \quad H_1: a\neq 0$ 的拒绝域为

$$R_1 \le C_U\left(\frac{\alpha}{2}\right) \text{ 或 } R_1 \ge C_L\left(\frac{\alpha}{2}\right).$$



第八节 假设检验问题的p值法

一、p值法

二、典型例题

三、小结

一、p值法

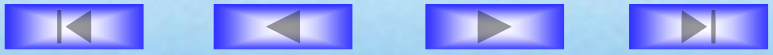
假设检验方法
- 临界值法.
- p值检验法

例1 设总体 $X \sim N(\mu,\sigma^2)$，μ 未知，$\sigma^2 = 100$，现有样本 $x_1, x_2, \cdots, x_{52}$，算得 $\bar{x} = 62.75$.

现在来检验假设

$$H_0: \mu \le \mu_0 = 60, \quad H_1: \mu > 60.$$

采用Z检验法,检验统计量为



$$z=\frac{\overline{X}-\mu_0}{\sigma/\sqrt{n}}.$$

以数据代入,得Z的观察值为

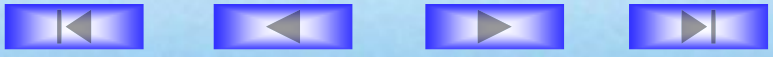
$$z_0=\frac{62.75-60}{10/\sqrt{52}}=1.983.$$

概率

$$P\{Z\geq z_0\}=P\{Z\geq 1.983\}=1-\Phi(1.983)=0.0238.$$

此即为图中标准正态曲线下位于z_0右边的尾部面积.

此概率称为Z检验法的右边检验的*p*值.



记为p值$=P\{Z \geq z_0\}=0.0238.$

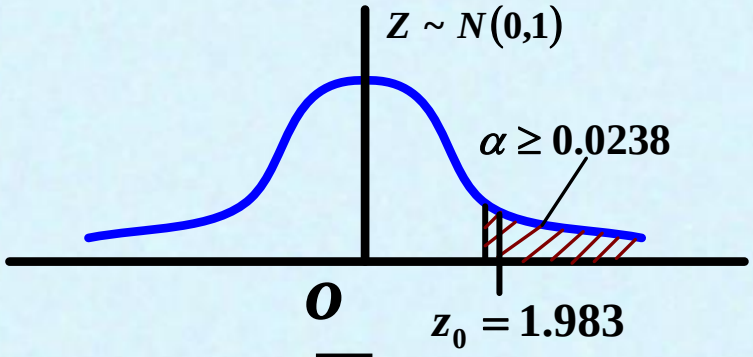


图1

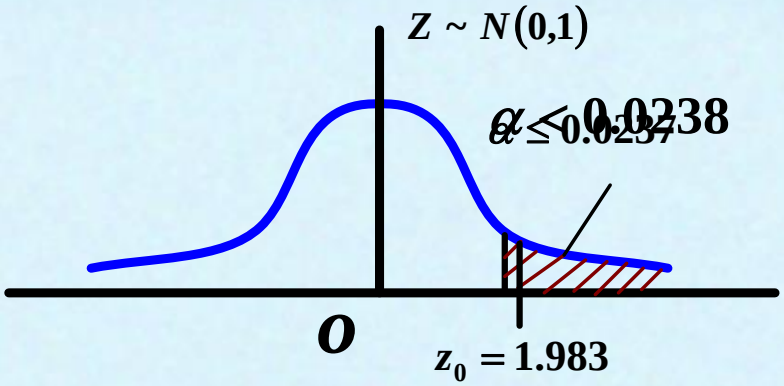


图2

若显著性水平$\alpha \geq p=0.0238$,则对应的临界值$z_\alpha \leq 1.983$,这表示观察值z_α=1.983落在拒绝域内 (如图1), 因而拒绝H_0; 又显著性水平$\alpha < p=0.0238$, 则对应的临界值$z_0 > 1.983$,这表示观察值 z_0=1.983 不落在拒绝域内图(2), 因而接受H_0.

定义 假设检验问题的*p*值(*probability value*)是由检验统计量的样本观察值得出的原假设可被拒绝的最小显著性水平.

任一检验问题的*p*值可以根据检验统计量的样本观察值的以及检验统计量在H_0下一个特定的参数值(一般是 H_0与H_1所规定的参数的分界点)对应的分布求出.

例如在正态分布$N(\mu,\sigma^2)$均值的检验中, 当σ未知时, 可采用检验统计量

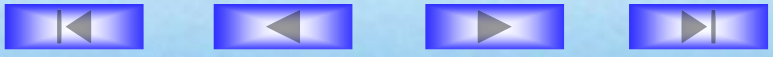
$t=\dfrac{\overline{X}-\mu_0}{S/\sqrt{n}}$,在以下三个检验问题中, 当$\mu=\mu_0$时, $t\sim t(n-1)$.如果由样本求得统计量t的观察值为t_0, 那么在检验问题

$$H_0:\mu\le\mu_0,\quad H_1:\mu>\mu_0\text{中}$$

p值$=P_{\mu_0}\{t\ge t_0\}=t_0$右侧尾部面积, 如图3;

$$H_0:\mu\ge\mu_0,\quad H_1:\mu<\mu_0\text{中}$$

p值$=P_{\mu_0}\{t\le t_0\}=t_0$左侧尾部面积, 如图4;



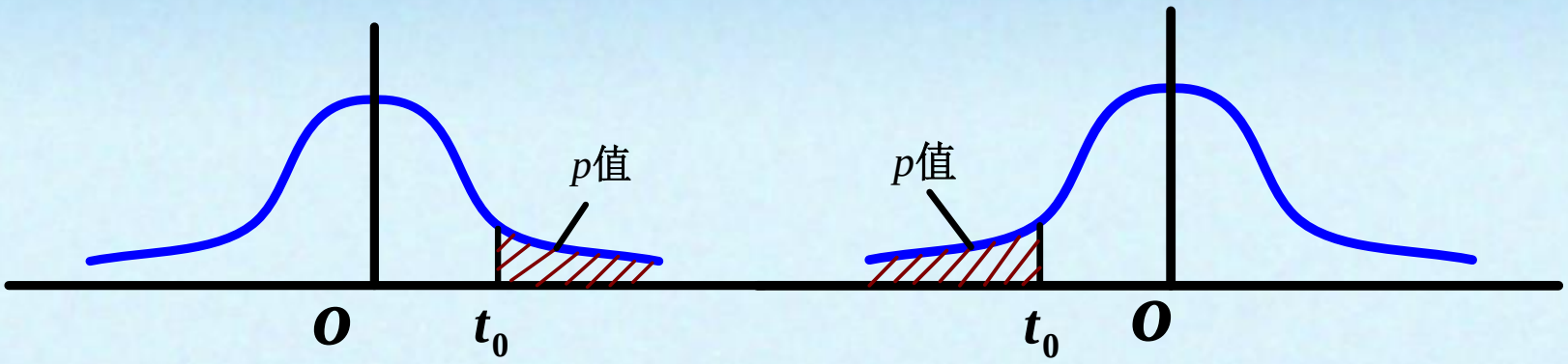


图3

图4

$H_0: \mu = \mu_0$，　$H_1: \mu \neq \mu_0$中

(*i*)当$t_0 > 0$时

$$p\text{值} = P_{\mu_0}\{|t| \geq t_0\} = P_{\mu_0}\{\{t \leq -t_0\} \cup \{t \geq t_0\}\}$$

$= 2\times(t_0$右侧尾部面积)如图5;

(*ii*)当$t_0 < 0$时

$$p\text{值} = P_{\mu_0}\{|t| \geq -t_0\} = P_{\mu_0}\{\{t \leq t_0\} \cup \{t \geq -t_0\}\}$$

综合$(i)(ii)$,

p值$= 2\times$(由t_0界定的尾部面积)如图6;

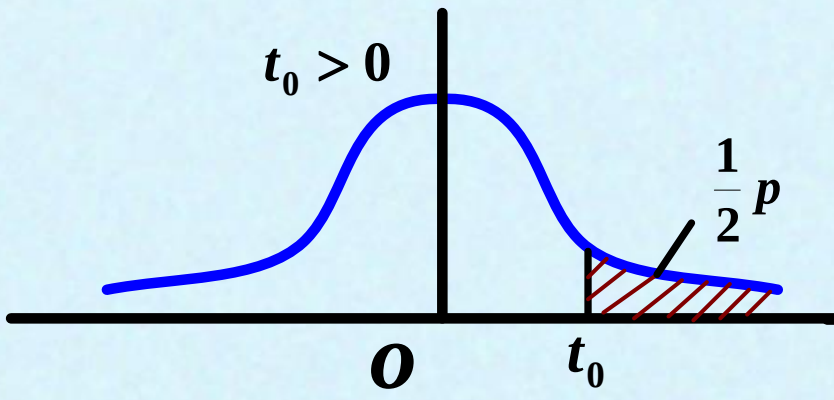


图5

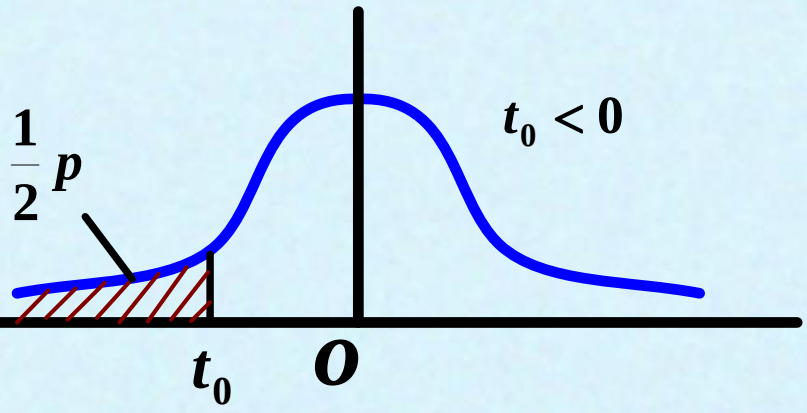


图6

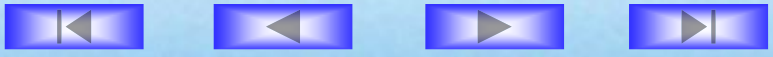
上述各图中的曲线均为$t(n-1)$分布的概率密度曲线.

在现代计算机统计软件中, 一般都给出检验问题的p值. 按p值的定义,对于任意指定的显著性水平α, 就有

⑴ 若p值$\le \alpha$, 则在显著性水平α下拒绝H_0;

⑵ 若p值$> \alpha$, 则在显著性水平α下接受H_0.

有了这两条结论就能方便地确定是否拒绝H_0. 这种利用p值来确定是否拒绝H_0的方法, 称为p值法.



用临界值法来确定H_0的拒绝域时，例如当$\alpha = 0.05$时知道要拒绝H_0，再取$\alpha = 0.01$也要拒绝H_0，但不能知道将α再降低一些是否也要拒绝H_0. 而p值法给出了拒绝H_0的最小显著性水平．因此p值法比临界值法给出了有关拒绝域的更多的信息.

二、典型例题

例2　用p值法检验本章第一节例2的检验问题

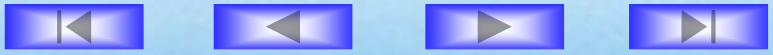
$$H_0: \mu \le \mu_0 = -0.545, \quad H_1: \mu > \mu_0 \quad \alpha = 0.05$$

解　用Z检验法，现在检验统计量$Z = \dfrac{\overline{x} - \mu_0}{\sigma/\sqrt{n}}$的观察值为

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

$$p\text{值} = P\{Z \ge 2.7955\} = 1 - \Phi(2.7955) = 0.0026.$$

p值$< \alpha = 0.05$，故拒绝H_0.



例 3 用p值法检验本章第二节例1的检验问题

$$H_0: \mu \le \mu_0 = 225,\ H_1: \mu > 225,\ \alpha = 0.05.$$

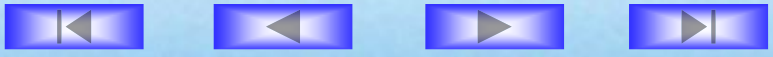
解 用t检验法，现在检验统计量$t = \dfrac{\overline{X} - \mu_0}{S/\sqrt{n}}$的观察值为

$$t = \frac{241.5 - 225}{98.7259/\sqrt{16}} = 0.6685.$$

由计算机算得

$$p\text{值} = P\{t \ge 0.6685\} = 0.2570.$$

$$p\text{值} > \alpha = 0.05,\ \text{故接受} H_0.$$



例 4　用p值法检验本章第三节例1的检验问题

$$H_0:\sigma^2=\sigma_0^{\,2}=5000,\ H_1:\sigma^2\neq 5000,\ \alpha=0.02.$$

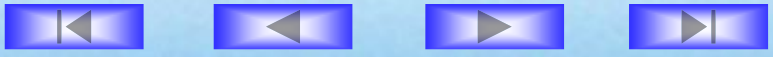
解　用χ^2检验法，现在检验统计量 $\chi^2=\dfrac{(n-1)S^2}{\sigma_0^2}$ 的观察值为

$$\chi_0^{\,2}=\frac{25\times 9200}{5000}=46.$$

由计算机算得

$$p\text{值}=2\times P\{\chi^2\geq 46\}=0.0128.$$

$$p\text{值}<\alpha=0.02,\ \text{故拒绝}H_0.$$



p值表示反对原假设H_0的依据的强度，p值越小，反对H_0的依据越强、越充分 (譬如对于某个检验问题的检验统计量的观察值的p值 $= 0.0009$，如此地小，以至于几乎不可能在 H_0为真时出现目前的观察值，这说明拒绝H_0的理由很强，我们就拒绝H_0.

一般, 若p值 ≤ 0.01, 称推断拒绝H_0的依据很强或称检验是高度显著的;

若$0.01 < p$值 ≤ 0.05, 称判断拒绝H_0的依据是强的或称检验是显著的;

若$0.05 < p$值 ≤ 0.1, 称推断拒绝H_0的理由是弱的, 检验是不显著的;

若p值 > 0.1, 一般来说没有理由拒绝.

基于p值, 研究者可以使用任意希望的显著性水平来作计算.

在杂志上或在一些技术报告中, 许多研究者在讲述假设检验的结果时, 常不明显地论及显著性水平以及临界值, 代之以简单地引用假设检验的 p值, 利用或让读者用它来评价反对原假设的依据的强度作出判断.

三、小结

假设检验方法
- 临界值法.
- p值法