第 七章 假设检验

- 1960-1980，在美国南方有好几起案件引起争议，有人指出陪审团的选择存在种族歧视，陪审团名单上的黑人很少

纯属巧合

- 证据表明：50%的合格公民是黑人，80人的预备陪审员名单上只有4名黑人

- 假设陪审员是随机选出的，80人的陪审团中黑人数量就应该服从二项分布B(80,0.5)

- 而最多有4名黑人的概率是

$$P(X \le 4) = 1.37\times10^{-18}$$

- 这是随机?

例. 体重指数BMI是目前国际上常用的衡量人体胖瘦程度以及是否健康的一个标准. 专家指出, 健康成年人的BMI取值应在18.55-24.99之间.

某种减肥药广告宣称, 连续使用该种减肥药一个星期便可达到减肥的效果.

为了检验其说法是否可靠,随机抽取9位试验者(要求BMI指数超过25、年龄在20-25岁女生),先让每位女生记录没有服用减肥药前的体重,然后让每位女生服用该减肥药,服药期间,要求每位女生保持正常的饮食习惯,连续服用该减肥药1周后,再次记录各自的体重.

测得服减肥药前后的体重差值(服药前体重-服药后体重)(单位:kg):

1.5, 0.6, -0.3, 1.1, -0.8, 0, 2.2, -1.0, 1.4

问题:根据目前的样本资料能否认为该减肥药广告中的宣称是可靠的?

什么是假设检验？

1. **概念**
    1. 事先对总体参数或分布形式作出某种假设
    2. 然后利用样本信息来判断原假设是否成立
2. **类型**
    1. 参数假设检验
    2. 分布假设检验
3. **特点**
    - 采用逻辑上的反证法
    - 依据统计上的小概率原理

原假设和备择假设

- **什么是原假设？(Null Hypothesis)**

待检验的假设，又称“0假设”

- 表示为 H_0
  - H_0： μ = 某一数值
  - 例如, H_0： $\mu = 3190$(克), $\mu \geq 3910$(克),或$\mu \leq 3910$(克)

- **什么是备择假设？(Alternative Hypothesis)**

  与原假设对立的假设

- 表示为 H_1
  - H_1： $\mu \neq$ 某一数值
  - 例如, H_1： $\mu \neq 3910$(克) , $\mu < 3910$(克),或$\mu > 3910$(克)

假设检验的过程

假设检验的过程是一个四步曲.

第一步，建立两个完全对立的假设:

原假设(零假设)H_0，备择假设(对立假设) H_1。

原假设与备择假设是不对称的！

决定谁是原假设，依赖于立场、惯例、方便性.

决定谁是原假设，依赖于立场、惯例、方便性.

1. 保护原假设.如果错误地拒绝假设A比错误地拒绝假设B带来更严重的后果——A选作原假设！

例如：假设A:新药有某种毒副作用，假设B:新药无某种毒副作用.

"有毒副作用"错误地当成"无毒副作用"比"无毒副作用"错误地当成"有毒副作用"带来的后果更严重。

2. 原假设为维持现状.为解释某些现象或效果的存在性，原假设常取为“无效果”、“无改进”、“无差异”等，拒绝原假设表示有较强的理由支持备择假设.

例1中原假设H_0:药物没有减肥效果.

备择假设H_1: 药物有减肥效果.

3. 原假设取简单假设.只有一个参数(或分布)的假设称为简单假设.如果只有一个假设是简单假设，将其取为原假设.

参数假设的形式

设θ是反映总体指标某方面特征的量，是我们感兴趣的参数. 一般参数θ的假设有三种情形:

H_0: $\theta=\theta_0$, H_1: $\theta<\theta_0$(左边检验)

H_0: $\theta=\theta_0$, H_1: $\theta>\theta_0$(右边检验)

H_0: $\theta=\theta_0$, H_1: $\theta\neq\theta_0$(双边检验)

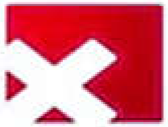
如何检验假设？

根据收集的资料，针对假设，给出检验方法，然后对假设进行判断。

判断方法有二种：临界值法．P_值法．

例　减肥药有效？

还是无效？



设服用减肥药前后体重差值$X \sim N(\mu, \sigma^2)$，
并假定方差 $\sigma^2 = 0.36$.

检验假设： $H_0 : \mu = 0, H_1 : \mu > 0,$

注意到：$\overline{X}$是μ的无偏估计，$\overline{X}$的取值大小反映了μ的取值大小，当原假设成立时，$\overline{X}$取值应偏小。因此

当$\overline{X} \geq C$时，拒绝原假设H_0，
当$\overline{X} < C$时，接受原假设H_0，
其中C是待定的常数.

如果统计量$T=T(X_1,\cdots,X_n)$的取值大小和原假设H_0是否成立有密切联系，可将其称为对应假设问题的**检验统计量**，而对应于拒绝原假设H_0时，样本值的范围称为**拒绝域**，记为W，其补集$\overline{W}$称为**接受域**.

第二步：给出检验统计量，并确定拒绝域的形式.

本例中的检验统计量为$\bar{X}$，拒绝域为

$$W=\{(X_1,\cdots,X_n):\bar{X}\geq C\}$$

- 如何选择C？ → 关键问题

一类、二类错误

由于样本的随机性，任一检验规则在应用时，都有可能发生错误的判断——两类错误.

	原假设为真	原假设不真
根据样本拒绝原假设	第I类错误	正确
根据样本接受原假设	正确	第II类错误

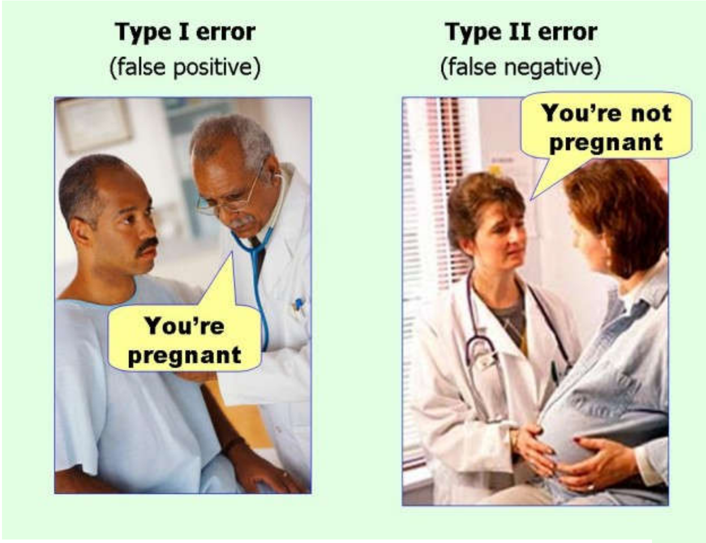
第I类错误：拒绝真实的原假设(弃真).

第II类错误：接受错误的原假设(取伪).

一类、二类错误

- 犯第一类错误的概率：
  P(拒绝H0|H0)=α

- 犯第二类错误的概率：
  P(接受H0|H1)



Neyman–Pearson原则：

首先控制犯第I类错误的概率不超过某个常数 $\alpha \in (0,1)$ ，再寻找检验，使得犯第II类错误的概率尽可能小. α称为显著水平.

常取 $\alpha = 0.01$, 0.05, 0.1等.

临界值法

第三步，根据显著水平和统计量的分布确定临界值——临界值法

取显著水平 $\alpha = 0.05$，

当 $H_0: \mu = 0$ 成立时，$\Rightarrow \dfrac{\bar{X}}{0.6/\sqrt{9}} \sim N(0,1)$，(统计量的分布)

犯第I类错误的概率可如下计算：

$$P\{\bar{X} \geq C \mid \mu = 0\} = P\left\{\frac{\bar{X}}{\sigma/\sqrt{n}} \geq \frac{C}{\sigma/\sqrt{n}} \middle| \mu = 0\right\}$$

$$= 1 - \Phi\left(\frac{C}{\sigma/\sqrt{n}}\right) \leq \alpha = 0.05. \quad (0.05 = \Phi(-z_{0.05}))$$

$$\Rightarrow \frac{C}{0.6/\sqrt{9}} \geq z_{0.05} = 1.645 \Rightarrow C \geq 0.329.$$

临界值法

根据Neyman－Pearson原则，为使犯第II类错误的概率尽可能小，应取$C=0.329$.因此，拒绝域$W=\{\bar{X}\geq 0.329\}$.

第四步：根据样本得出结论.

根据实际样本资料，得$\bar{x}=0.522>0.329$.

当原假设H_0成立时，样本落在拒绝域的概率不超过0.05，是小概率事件。

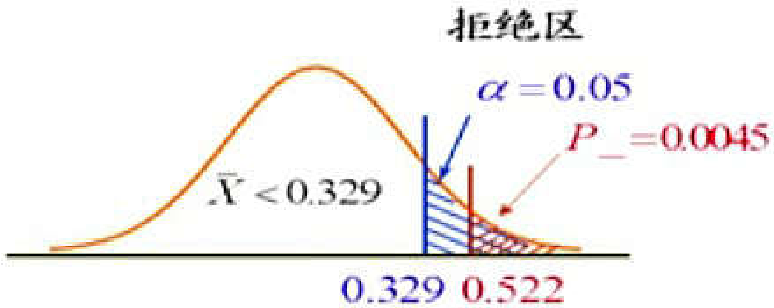
根据实际推断原理，有充分的理由拒绝原假设，认为厂家的宣传是可靠的.

同理，若$\alpha=0.01$,拒绝域$W=\{\bar{X}\geq 0.465\}$，拒绝原假设.

P值法

第三’ 步：计算最小显著水平——P_值法

P_-值： 当原假设H_0成立时，检验统计量取比观察到的结果更为极端的数值的概率.



$$P_-=P\{\bar{X}\geq\bar{x}=0.522\,|\,\mu=0\}$$

$$=1-\Phi\left(\frac{0.522}{0.6/\sqrt{9}}\right)=0.0045<\alpha=0.05$$

概率这么小的事件！
竟然发生了！！
拒绝原假设！！！

第四’ 步：比较P_值与显著水平，得出结论.

P值法

P 值与显著水平 α 的关系：

(1) 若 $P \le \alpha$，等价于样本落在拒绝域内，因此，拒绝原假设，称检验结果在水平 α 下是统计显著的.

(2) 若 $P > \alpha$，等价于样本不落在拒绝域内，因此，不拒绝（接受）原假设，称检验结果在水平 α 下是统计不显著.

假设检验的步骤

①根据题意或预期反向提出原假设H0、备择假设H1

②选择检验统计量，其分布不依赖于任何参数

③在给定显著性水平α下，确定拒绝域

④根据实际样本观测值做出判断

临界值法

①根据题意或预期反向提出原假设H0、备择假设H1

②选择检验统计量，其分布不依赖于任何参数

③计算检验统计量的观测值与P值

④根据给定的显著水平α做出判断

P值法

假设检验

- 对于正态总体$N(\mu,\sigma^2)$，σ^2已知，$X_1,\ldots,X_n$为来自正态总体的样本，可用$U=\frac{\bar{X}-\mu}{\sigma/\sqrt{n}}$作为检验统计量

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

- 拒绝域的形式为

$$|u|=\left|\frac{\bar{x}-\mu_0}{\sigma/\sqrt{n}}\right|\geq k\ (k\text{待定})$$

假设检验

- 犯第一类错误的概率为

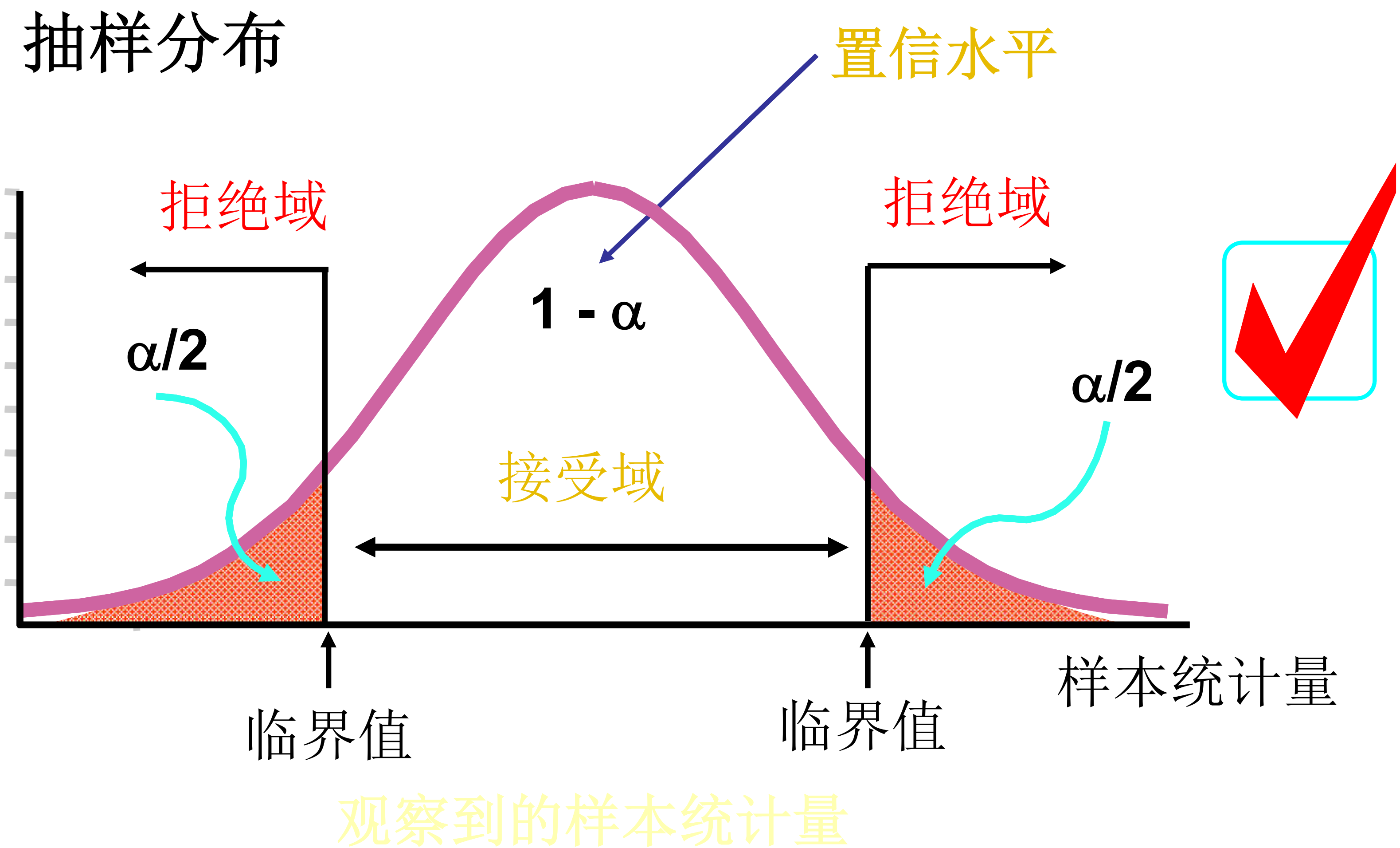
$$P(\text{拒绝H0})|\mu=\mu_0)=P(|u|=|\frac{\bar{x}-\mu_0}{\frac{\sigma}{\sqrt{n}}}|\geq k|\mu=\mu_0)$$

- 给定显著水平α，查表得k=$u_{\alpha/2}$,故拒绝域为

$$|u|=|\frac{\bar{x}-\mu_0}{\frac{\sigma}{\sqrt{n}}}|\geq u_{\alpha/2}$$

双边检验

（显著性水平与拒绝域）



- 对于

$$H_0\colon \mu = \mu_0, H_1\colon \mu > \mu_0,$$

- 犯第一类错误的概率为

$$P(\text{拒绝H0})|\mu = \mu_0) = P(u = \frac{\bar{x}-\mu_0}{\frac{\sigma}{\sqrt{n}}} \geq k|\mu = \mu_0)$$

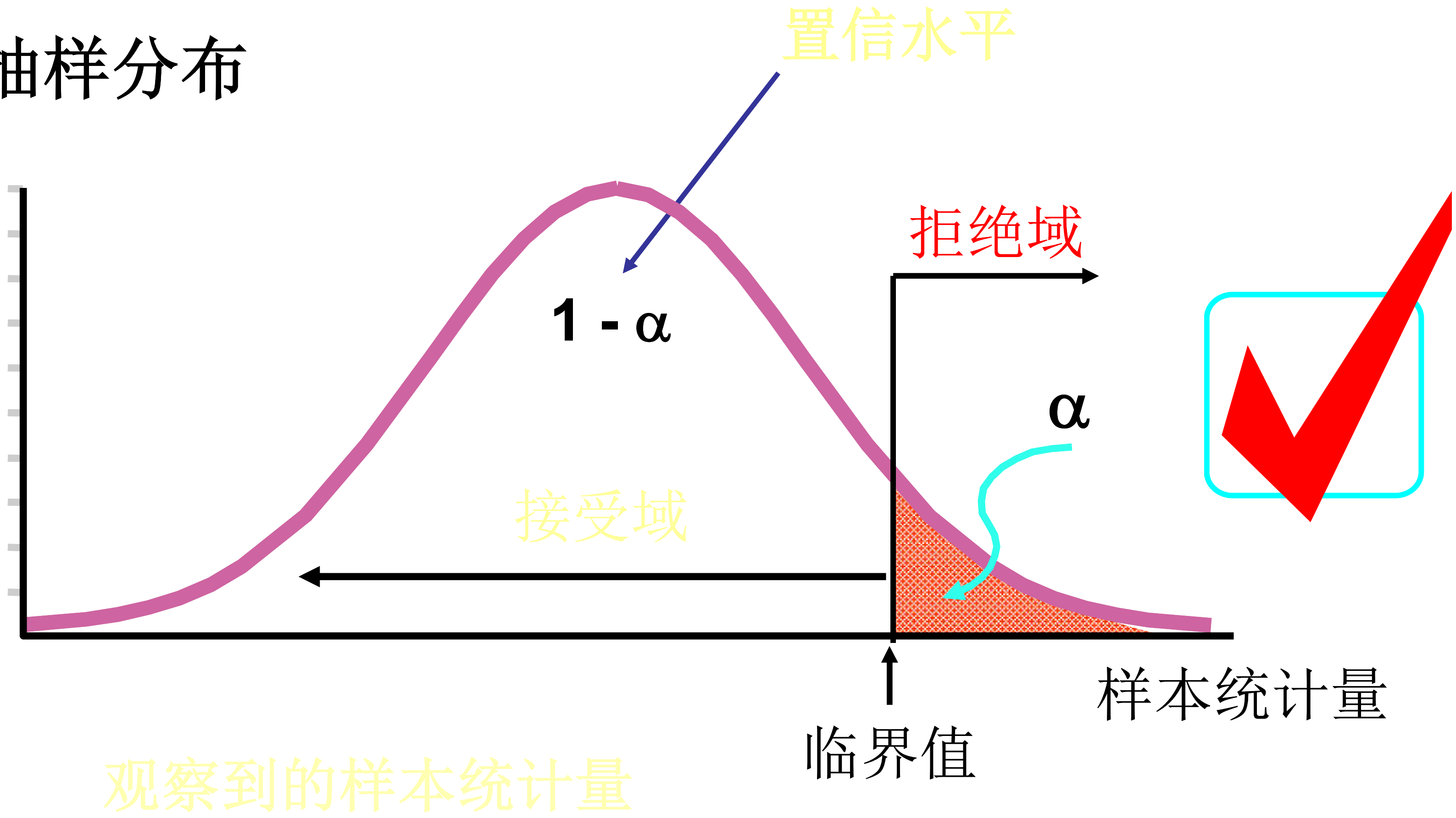
- 于是拒绝域为

$$u = \frac{\bar{x} - \mu_0}{\frac{\sigma}{\sqrt{n}}} \geq u_\alpha$$

右边检验

（显著性水平与拒绝域 ）

抽样分布



假设检验

- 对于

$$H_0:\ \mu = \mu_0, H_1:\ \mu < \mu_0,$$

- 犯第一类错误的概率为

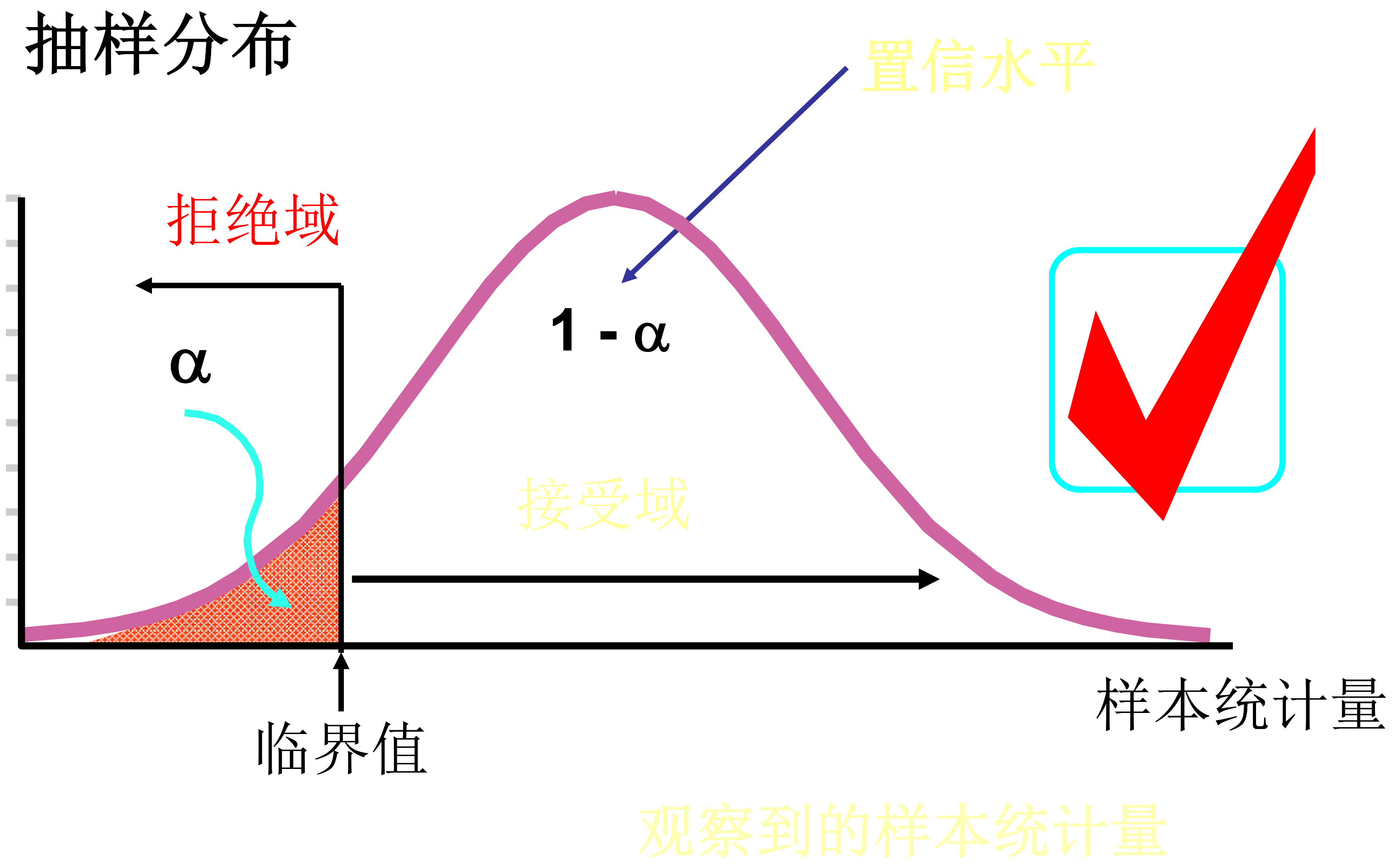
$$P(\text{拒绝H0})|\mu = \mu_0) = P(u = \frac{\bar{x}-\mu_0}{\frac{\sigma}{\sqrt{n}}} \le k|\mu = \mu_0)$$

- 于是拒绝域为

$$u = \frac{\bar{x} - \mu_0}{\frac{\sigma}{\sqrt{n}}} \le -u_\alpha$$

左边检验

（显著性水平与拒绝域 ）



均值的双边U检验

1. 假定条件

- 总体服从正态分布，(σ^2 已知)
- 若不服从正态分布，可用正态分布来近似($n\geq30$)

2. 原假设: H_0: $\mu=\mu_0$， 备择假设: H_1：$\mu \neq \mu_0$

3.使用U检验法

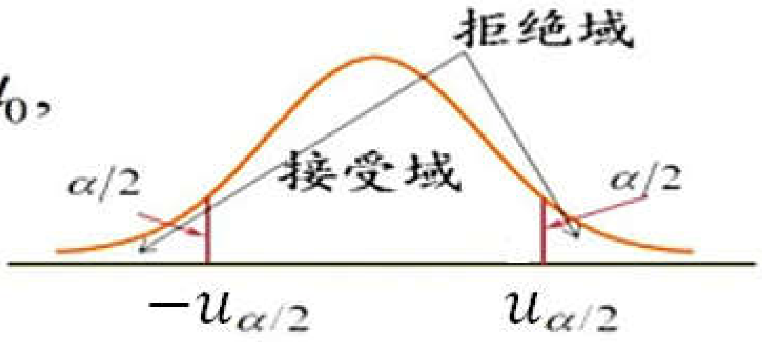
$$U=\frac{\bar{X}-\mu_0}{\sigma/\sqrt{n}}\sim N(0,1)$$

均值的双边U检验

双边假设问题

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

其中μ_0是已知的常数.



检验统计量为 $U = \dfrac{\bar{X} - \mu_0}{\sigma/\sqrt{n}}$

检验拒绝域$W = \left\{ |U| = \left| \dfrac{\bar{X} - \mu_0}{\sigma/\sqrt{n}} \right| \geq u_{\alpha/2} \right\}.$

均值的双边U检验

P_值的计算

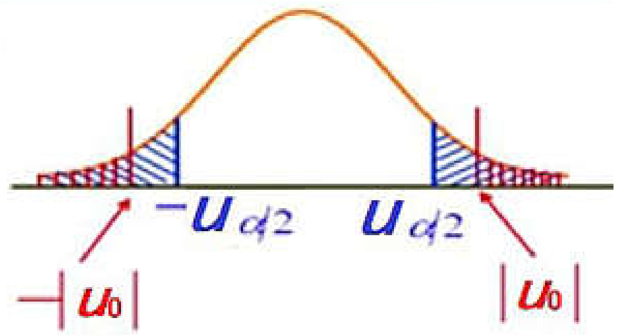
对给定的样本观察值$x_1,\cdots,x_n$，记检验统计量U的取值

$$u_0=\frac{\bar{x}-\mu_0}{\sigma/\sqrt{n}},$$

$$P\_=P_{H_0}\{|U|\geq|u_0|\}=2(1-\Phi(|u_0|)).$$

当$P_\leq\alpha$时，拒绝原假设，

当$P_>\alpha$时，接受原假设.



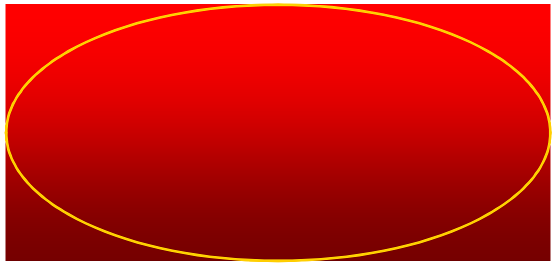
红色区域概率值：$P_$值

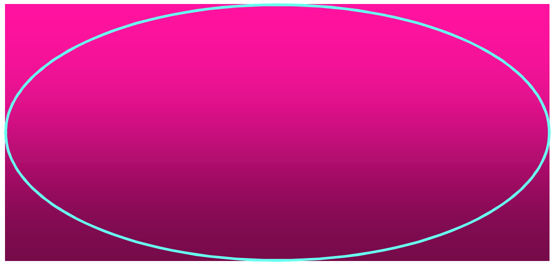
蓝色区域概率值：α

$P_$值$<\alpha$，拒绝H_0.

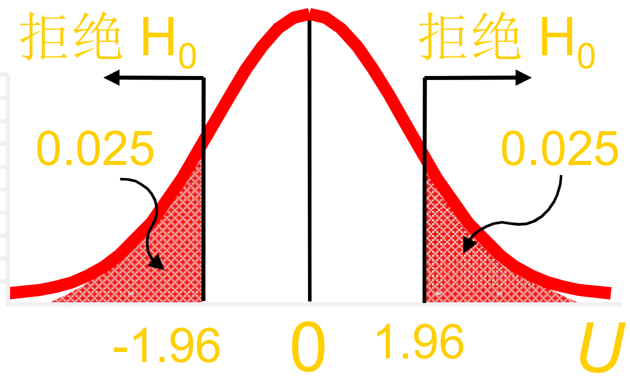
均值的双边*U*检验

- 【例】某机床厂加工一种零件，根据经验知道，该厂加工零件的椭圆度近似服从正态分布，其总体均值为μ_0=0.081mm，总体标准差为σ= 0.025 。今换一种新机床进行加工，抽取n=200个零件进行检验，得到的平均椭圆度为0.076mm。试问新机床加工零件的椭圆度的均值与以前有无显著差异？（α＝0.05）





- **H_0: μ = 0.081**
- **H_1: $\mu \neq$ 0.081**
- **σ= 0.025**
- **α = 0.05**
- **n = 200**
- **临界值(s):**



检验统计量:

$$u = \frac{\bar{x} - \mu_0}{\sigma/\sqrt{n}} = \frac{0.076 - 0.081}{0.025/\sqrt{200}} = -2.83$$

决策: $|u| > |u_{\alpha/2}| = |u_{0.025}| = 1.96$

拒绝H_0

结论:

有证据表明新机床加工的零件的椭圆度与以前有显著差异

例： 据健康统计中心报告35至44岁的男子心脏收缩压服从正态分布,平均心脏收缩压为128，标准差为15. 现根据某公司在35至44岁年龄段的72位员工的体检记录，计算得平均心脏收缩压为126. 07(mm/hg)． 问该公司员工的心脏收缩压与一般人群是否存在差异呢?（假设该公司员工的心脏收缩压与一般中年男子的心脏收缩压具有相同的标准差）。（α=0.05）

步骤1：提出检验假设

$$H_0 : \mu = 128, H_1 : \mu \neq 128$$

步骤2:计算检验统计量的观测值.

$$u_0 = \frac{\overline{x} - \mu_0}{\sigma / \sqrt{n}} = \frac{126.07 - 128}{15 / \sqrt{72}} = -1.09$$

步骤3：计算P_值

$$P_{-}=2(1-\Phi(|u_0|))=2(1-\Phi(1.09))=0.2758.$$

步骤4：根据实际情况作出判断

P_=0.2758>0.05，因此，没有充分理由拒绝原假设。

均值的单边*U*检验

1. 假定条件

- 总体服从正态分布，(σ^2 已知)
- 若不服从正态分布，可用正态分布来近似($n \geq 30$)

2. 原假设: H_0: $\mu=\mu_0$，　备择假设: H_1 : $\mu > \mu_0$

　　H_0: $\mu=\mu_0$,　　　　　　H_1 : $\mu < \mu_0$

3.使用*U*检验法

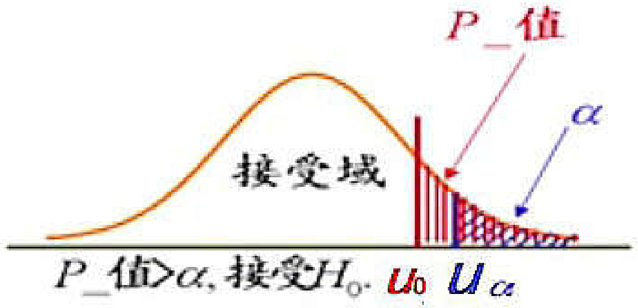
$$U = \frac{\bar{X} - \mu_0}{\sigma/\sqrt{n}} \sim N(0,1)$$

西安交通大学 XI'AN JIAOTONG UNIVERSITY

均值的单边U检验

右边假设问题：$H_0:\mu=\mu_0, H_1:\mu>\mu_0$，

其中μ_0是已知的常数.



检验统计量为 $U=\dfrac{\overline{X}-\mu_0}{\sigma/\sqrt{n}}$

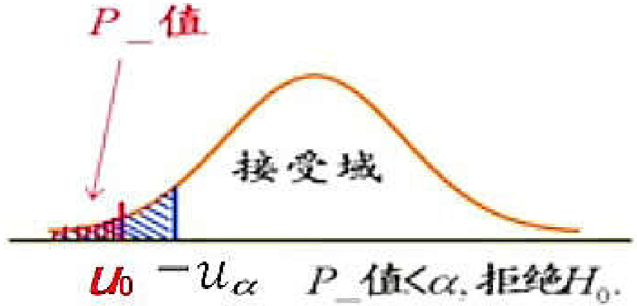
检验拒绝域$W=\left\{U=\dfrac{\overline{X}-\mu_0}{\sigma/\sqrt{n}}\geq u_\alpha\right\}$.

$P\_=P_{H_0}\{U\geq u_0\}=1-\Phi(u_0)$. 其中$u_0=\dfrac{\overline{x}-\mu_0}{\sigma/\sqrt{n}}$.

均值的单边U检验

左边假设问题：$H_0: \mu = \mu_0, H_1: \mu < \mu_0$，

其中μ_0是已知的常数.



检验统计量为 $U = \dfrac{\bar{X} - \mu_0}{\sigma/\sqrt{n}}$

检验拒绝域$W = \left\{ U = \dfrac{\bar{X} - \mu_0}{\sigma/\sqrt{n}} \le -u_\alpha \right\}$.

$P\_ = P_{H_0}\{U \le u_0\} = \Phi(u_0)$. 其中$u_0 = \dfrac{\bar{x} - \mu_0}{\sigma/\sqrt{n}}$.

西安交通大學 XI'AN JIAOTONG UNIVERSITY

均值的单边U检验

例：　为了了解A高校学生的消费水平，随机抽取225位学生调查其月消费(近6个月的消费平均值)，得到该225位学生的平均月消费为1530元。假设学生月消费服从正态分布，标准差为$\sigma=120$.

已知B高校学生的月平均消费为1550 元，是否可以认为A高校学生的消费水平要低于B高校？

均值的单边U检验

步骤1：提出检验假设

$$H_0: \mu = 1550, H_1: \mu < 1550$$

步骤2：确定检验规则

检验统计量为 $U = \dfrac{\bar{X} - 1550}{\sigma / \sqrt{n}}$. 取显著水平$\alpha = 0.05$,

由备择假设的形式知，这是左边检验，因此检验规则为：当$U \le -u_\alpha = -u_{0.05} = -1.645$时，拒绝$H_0$.

均值的单边U检验

步骤3: 计算检验统计量的值

将样本均值$\bar{x}=1530, \sigma=120, n=225,$

代入检验统计量，计算得

$$U=\frac{\bar{X}-1550}{\sigma/\sqrt{n}}=\frac{1530-1550}{120/\sqrt{225}}=-2.5<-1.645.$$

步骤4: 根据实际情况作出判断

因此，根据检验规则，做出拒绝原假设H_0的判断.

即认为A高校学生的生活水平低于B高校.

均值的单边U检验

利用P_值进行假设检验

步骤3' ：计算P_值

$$P = P\left(\frac{\bar{X}-1550}{\sigma/\sqrt{n}} \le \frac{1530-1550}{120/\sqrt{225}}\right)$$

$$= P(U \le -2.5) = 0.006$$

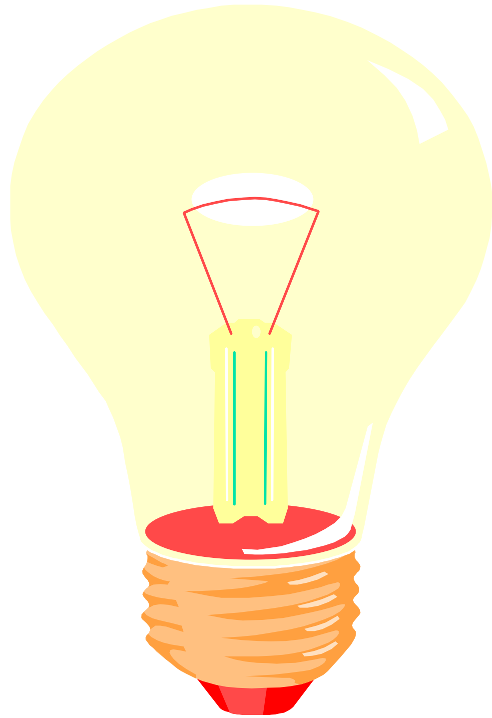
步骤4' ：根据显著水平作出判断

$$P = 0.006 < \alpha = 0.05,$$

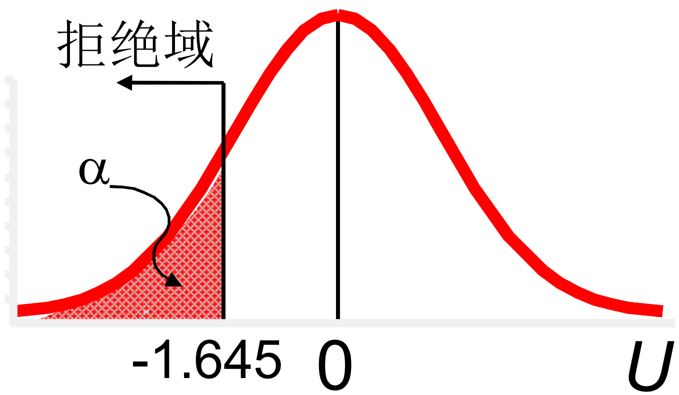
同样做出拒绝原假设$H_0 : \mu = 1550$的判断.

均值的单尾检验

- 【例】某批发商欲从生产厂家购进一批灯泡，根据合同规定，灯泡的使用寿命平均不能低于1000小时。已知灯泡使用寿命服从正态分布，标准差为200小时。在总体中随机抽取100只灯泡，测得样本均值为960小时。批发商是否应该购买这批灯泡？($\alpha＝0.05$)



- **H0: μ ≥ 1000**
- **H1: μ < 1000**
- σ= 200
- α = **0.05**
- n = **100**
- **临界值(s):**



检验统计量:

$$u = \frac{\bar{x} - \mu_0}{\sigma/\sqrt{n}} = \frac{960 - 1000}{200/\sqrt{100}} = -2$$

决策: $u < -u_\alpha = -u_{0.05} = -1.645$

在 α = 0.05的水平上拒绝H_0

结论:

有证据表明这批灯泡的使用寿命显著不满足大于等于1000小时的寿命条件

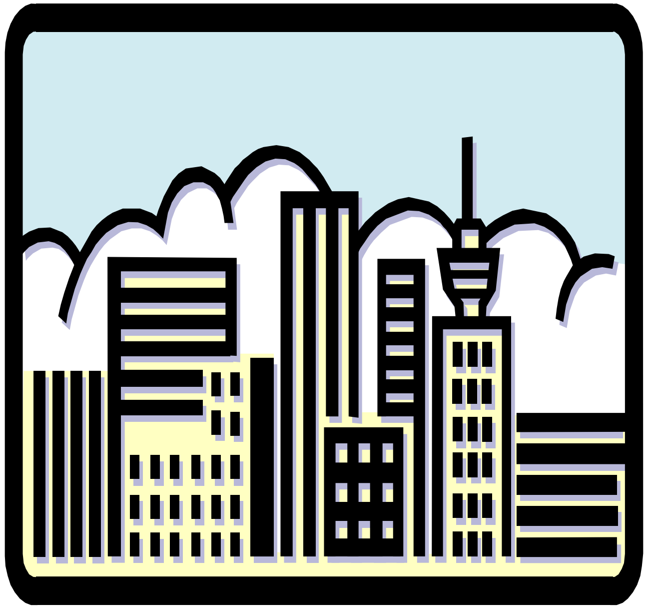
均值的双尾 t 检验

1. 假定条件
   - 总体为正态分布(σ^2 未知)
   - 若不服从正态分布，可以用正态分布来近似 ($n\geq30$)
2. 使用t检验法

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

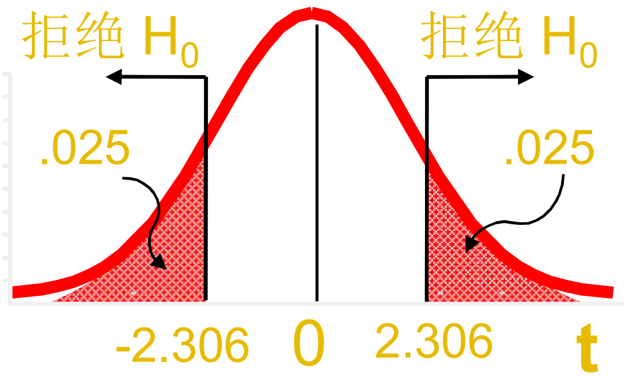
均值的双尾 t 检验

- 【例】某厂采用自动包装机分装产品，假定每包产品的重量服从正态分布，每包标准重量为1000克。某日随机抽查9包，测得样本平均重量为986克，样本标准差为24克。试问在0.05的显著性水平上，能否认为这天自动包装机工作正常？



均值的双尾 t 检验

- **H_0: μ = 1000**
- **H_1: $\mu \neq$ 1000**
- **α = 0.05**
- **df = 9 - 1 = 8**
- **临界值(s):**



检验统计量:

$$t = \frac{\bar{x} - \mu_0}{s/\sqrt{n}} = \frac{986 - 1000}{24/\sqrt{9}} = -1.75$$

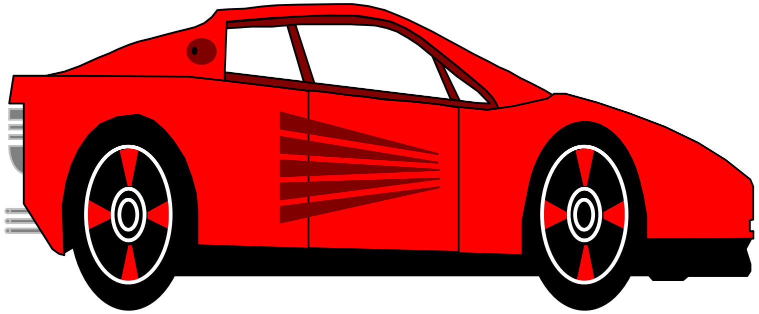
决策: $|t| = 1.75 < t_{\frac{\alpha}{2}}(8)$

在 α = 0.05的水平上接受H_0

结论:

有证据表明这天自动包装机工作正常

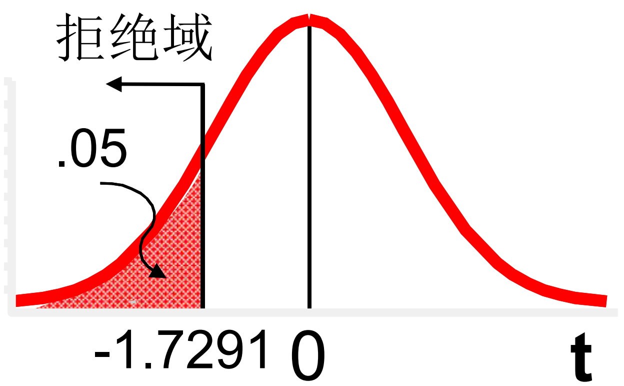
均值的单尾 t 检验



- 【例】一个汽车轮胎制造商声称，某一等级的轮胎的平均寿命在一定的汽车重量和正常行驶条件下大于40000公里，对一个由20个轮胎组成的随机样本作了试验，测得平均值为41000公里，标准差为5000公里。已知轮胎寿命的公里数服从正态分布，我们能否根据这些数据作出结论，该制造商的产品同他所说的标准相符？($\alpha = 0.05$)

均值的单尾 t 检验

- **H_0: μ ≥ 40000**
- **H_1: μ < 40000**
- **α = 0.05**
- **df = 20 - 1 = 19**
- **临界值(s):**



检验统计量:

$$t = \frac{\bar{x} - \mu_0}{s/\sqrt{n}} = \frac{41000 - 40000}{5000/\sqrt{20}} = 0.894$$

决策: $t = 0.894 > -t_{\alpha}(19)$

在α = 0.05的水平上接受H_0

结论:

有证据表明轮胎使用寿命显著地大于40000公里

方差的卡方 (χ^2) 检验

- 1. 假设总体近似服从正态分布(μ未知)
- 2. 检验一个总体的方差
- 3. 原假设为 H_0: $\sigma^2 = \sigma_0{}^2$
- 4. 检验统计量

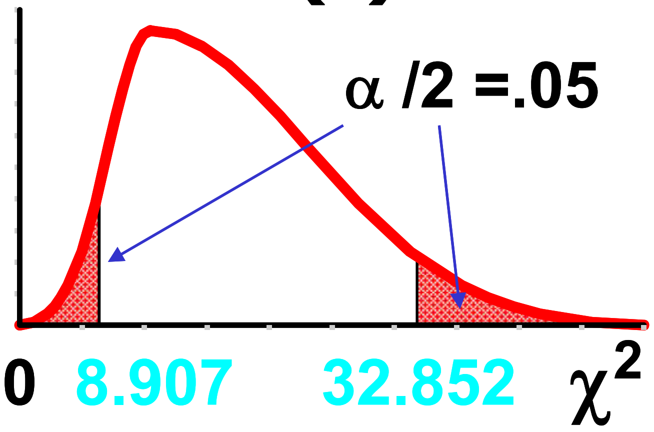
$$\chi^2 = \frac{(n-1)S^2}{\sigma_0^2} \sim \chi^2(n-1)$$

卡方 (χ^2)检验

- 【例】根据长期正常生产的资料可知，某厂所产维尼纶的纤度服从正态分布，其方差为**0.0025**。现从某日产品中随机抽取**20**根，测得样本方差为**0.0042**。试判断该日纤度的波动与平日有无显著差异？(α=**0.05**)

卡方 (χ^2) 检验

- **H_0: σ^2 = 0.0025**
- **H_1: $\sigma^2 \neq$ 0.0025**
- **α = 0.05**
- **df = 20 - 1 = 19**
- **临界值(s):**



统计量:

$$\chi^2 = \frac{(n-1)s^2}{\sigma_0^2}$$
$$= \frac{(20-1)0.0042}{0.0025}$$
$$= 31.92$$

决策: $\chi^2 < \chi^2_{\frac{\alpha}{2}}(19)$

在 α = 0.05的水平上接受H_0

结论:

有证据表明该日纤度的波动比平时没有显著差异

两个总体均值之差的 t 检验

1. 检验具有等方差的两个总体的均值(σ_1^2、σ_2^2未知)
2. 假定条件
   - 两个总体都是正态分布
   - 两个总体方差未知但相等$\sigma_1^2=\sigma_2^2$
3. 原假设为H_0: $\mu_1 - \mu_2 = c$
4. t检验法

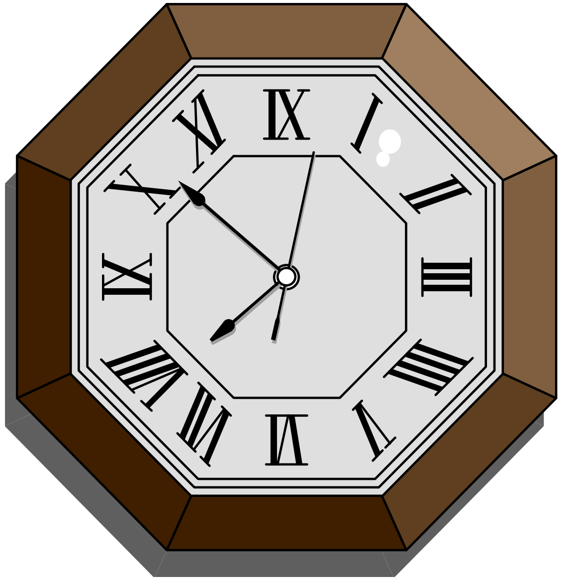
$$T = \frac{(\bar{X}_1 - \bar{X}_2) - c}{S_w\sqrt{\frac{1}{n_1}+\frac{1}{n_2}}} \sim t(n_1+n_2-2)$$

其中

$$S_w^2 = \frac{(n_1-1)S_{1n_1}^2 - (n_2-1)S_{2n_2}^2}{n_1+n_2-2}$$

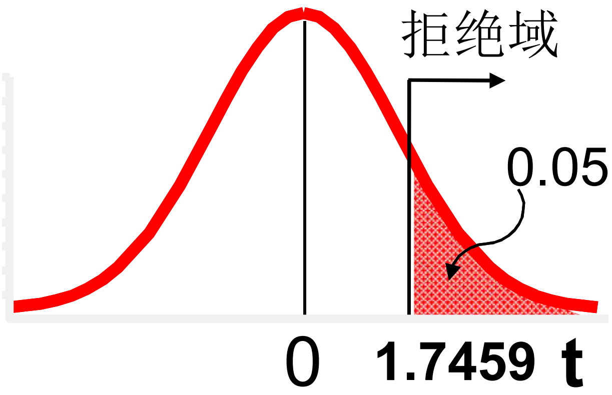
两个总体均值之差的 t 检验

【例】一个车间研究用两种不同的工艺组装某种产品所用的时间是否相同。让一个组的**10**名工人用第一种工艺组装该产品，平均所需时间为**26.1**分钟，样本标准差为**12**分钟；另一组**8**名工人用第二种工艺组装，平均所需时间为**17.6**分钟，样本标准差为**10.5**分钟。已知用两种工艺组装产品所用时间服从正态分布，且$\sigma_1^2=\sigma_2^2$。试问能否认为用第二种方法组装比用第一中方法组装更好？**(α = 0.05)**



两个总体均值之差的 t 检验

- **H0: $\mu_1 - \mu_2 \leq 0$**
- **H1: $\mu_1 - \mu_2 > 0$**
- **$\alpha = 0.05$**
- **$n_1 = 10$，$n_2 = 8$**
- **临界值(s):**



检验统计量:

$$t = \frac{(\bar{x}_1 - \bar{x}_2) - 0}{s_w\sqrt{\frac{1}{n_1} + \frac{1}{n_2}}} = \frac{26.1 - 17.6 - 0}{11.37\sqrt{\frac{1}{10} + \frac{1}{8}}} = 1.576$$

决策: $t < t_\alpha(10 + 8 - 2)$

接受H_0

结论:

没有证据表明用第二种方法组装更好

两个总体方差之比的 F 检验

1. 检验两个总体的方差之比
2. 假定条件
   - 两个总体都是正态分布
   - 两个总体均值未知
   - 两个样本是独立的随机样本
3. 原假设为 $H_0: \frac{\sigma_1^2}{\sigma_2^2} = C$
4. 检验统计量 $F = \frac{S_{1n_1}^2}{CS_{2n_2}^2} \sim F(n_1 - 1, n_2 - 1)$

- 甲乙两厂生产同一种电阻，现从两厂的产品中分别抽取12和10个样品，测量它们的电阻值，计算出样本方差分别为S_1^2 =1.4，S_2^2 =4.38.假设电阻值服从正态分布，在显著性水平0.05下，能否认为两厂生产的电阻值方差相等？

$$H_0\colon \frac{\sigma_1^2}{\sigma_2^2} = 1 \quad H_1\colon \frac{\sigma_1^2}{\sigma_2^2} \neq 1$$

$$F = \frac{S_{1n_1}^2}{S_{2n_2}^2} = \frac{1.4}{4.38} = 0.32$$

$$F_{\alpha/2}(11,9) = 3.91$$

$$F_{1-\alpha/2}(11,9) = 1/F_{\alpha/2}(9,11) = 0.28$$

决策：接受H_0

成对数据的假设检验

成对数据t检验

配对研究的数据是一对一对地收集得到的，所以也称为成对数据的研究．由于配对研究采用了比较的思想，比通常的单个样本推断更让人信服．这种方法在医学和生物研究领域中广泛存在，成对数据检验的基本思想是将两样本问题转为单样本问题．

成对数据的假设检验

- 假设成对数据 $(X_1, Y_1), \cdots, (X_n, Y_n)$
- 设差值 $Z_i = X_i - Y_i, i = 1, \cdots, n.$
- 差值可以看成来自正态总体$N(\mu, \sigma^2)$的样本

可考虑如下的检验问题：

$$H_0: \mu = 0, H_1: \mu \neq 0$$

成对数据的假设检验

记 $\bar{Z}=\frac{1}{n}\sum_{i=1}^{n}Z_i$， $S_Z^2=\frac{1}{n-1}\sum_{i=1}^{n}(Z_i-\bar{Z})^2$，

则检验统计量为 $T=\frac{\bar{Z}-z_0}{S_Z/\sqrt{n}}$

检验的拒绝域为$W=\{|T|\geq t_{\alpha/2}(n-1)\}$，

观察值为 $t_0=\frac{\bar{z}-z_0}{s_Z/\sqrt{n}}$

P_值为：$P\_=P_{H_0}\{|T|\geq|t_0|\}=2P\{t(n-1)\geq|t_0|\}$.

配对样本的 t 检验

- 【例】一个以减肥为主要目标的健美俱乐部声称，参加其训练班至少可以使减肥者平均体重减重8.5公斤以上。为了验证该宣称是否可信，调查人员随机抽取了10名参加者，得到他们的体重记录如下表：

训练前	94.5	101	110	103.5	97	88.5	96.5	101	104	116.5
训练后	85	89.5	101.5	96	86	80.5	87	93.5	93	102

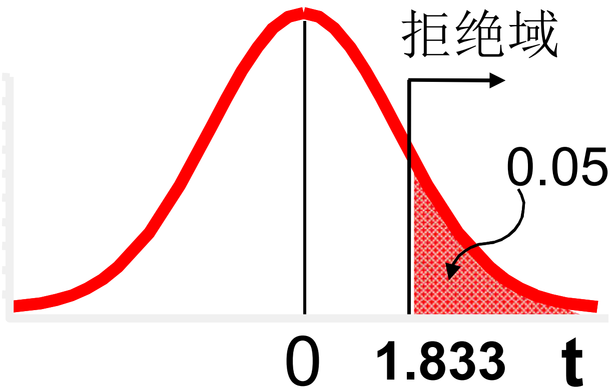
在 $\alpha = 0.05$的显著性水平下，调查结果是否支持该俱乐部的声称？

配对样本的 t 检验

样本差值计算表		
训练前	训练后	差值Z_i
94.5	85	9.5
101	89.5	11.5
110	101.5	8.5
103.5	96	7.5
97	86	11
88.5	80.5	8
96.5	87	9.5
101	93.5	7.5
104	93	11
116.5	102	14.5
合计	—	98.5

配对样本的 t 检验

- **H0: μ ≤8.5**
- **H1: μ > 8.5**
- **α = 0.05**
- **df = 10 - 1 = 9**
- **临界值(s):**



检验统计量:

$$t = \frac{\bar{z} - z_0}{s/\sqrt{n}} = \frac{9.85 - 8.5}{2.199/\sqrt{10}} = 1.94$$

决策: $t = 1.94 > t_{\alpha}(9) = 1.833$

拒绝H_0

结论:

有证据表明该俱乐部的宣称是可信的

皮尔逊定理

设一个随机试验的r个结果$A_1, A_2, \ldots, A_r$构成互斥完备事件群，在一次试验中它们发生的概率分别为$p_1, p_2, \ldots, p_r$，其中$p_i > 0(i = 1,2,\ldots,r)$，且$\sum_{i=1}^{r} p_i = 1$，以$m_i$表示$n$次独立重复试验中$A_i$发生的次数，那么，当$n \to \infty$，随机变量

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

的分布收敛于自由度为r-1的χ^2分布。

分布的假设检验

χ^2拟合检验法

设$X_1, \dots, X_n$是来自总体X的样本，$F(x)$是一个完全已知的分布函数，在显著水平α下，检验假设

$$H_0: X\text{的分布函数是}F(x)\text{，}H_1: X\text{的分布函数不是}F(x)$$

1. 将总体X的取值范围划分为r个互不相交的子集(区间)$A_1, \dots, A_r$，则其构成互斥完备事件群;
2. 在H_0下，计算事件A_i的概率$P(A_i) = p_i$，则事件A_i发生的频数为np_i;

χ^2拟合检验法

3. 统计出样本值$x_1,\dots,x_n$中，事件A_i发生的实际频数m_i;

4. 在显著水平α下，考虑统计量

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

的拒绝域

$$\chi^2=\sum_{i=1}^{r}\frac{(m_i-np_i)^2}{np_i}\geq\chi^2{}_{\alpha}\,(r-1)$$

注意：

- 样本容量较大，一般$n \geq 50$;
- 分组时每个区间所含样本数m_i不应少于5;
- 当$F(x)$是已知函数形式，但是其中含有l个未知参数$\theta_1, \ldots, \theta_l$时，一般先用极大似然估计法得到估计值$\hat{\theta}_1, \ldots, \hat{\theta}_l$，然后用$F\ (x; \hat{\theta}_1, \ldots, \hat{\theta}_l)$代替$F(x)$进行检验，拒绝域为

$$\chi^2 = \sum_{i=1}^{r} \frac{(m_i - np_i)^2}{np_i} \geq {\chi^2}_\alpha\ (r - l - 1)$$

例：孟德尔遗传理论断言，当两个品种的豆杂交时，圆的和黄的、起皱的和黄的、圆的和绿的、起皱的和绿的豆的频数将以比例9：3：3：1发生。在检验这个理论时，孟德尔分别得到频数315、101、108、32、这些数据是否支持该理论？

解：定义
$$X=\begin{cases}1, & \text{若豆子是圆的和黄的}\\ 2, & \text{若豆子是起皱的和黄的}\\ 3, & \text{若豆子是圆的和绿的}\\ 4, & \text{若豆子是起皱的和绿的}\end{cases}$$

$$H_0: p_1=P(X=1)=\frac{9}{16}, p_2=P(X=2)=\frac{3}{16},$$

$$p_3=P(X=3)=\frac{3}{16}, p_4=P(X=4)=\frac{1}{16}.$$

豆子状态 x	1	2	3	4
实测频数 m_i	315	101	108	32
概率 p_i	9/16	3/16	3/16	1/16
理论频数 np_i	312.75	104.25	104.25	34.75

$$\chi^2 = \sum_{i=1}^{4} \frac{(m_i - np_i)^2}{np_i} = 0.47 < \chi^2_{0.05}(3) = 7.815,$$

不拒绝原接受，即数据支持该理论.