大学数学实验

实验12 统计推断

（statistical inference）

清 华 大 学 数 学 科 学 系

实验12的基本内容

统计推断

一、实例及其分析

二、参数估计

三、假设检验

四、实例的求解

报道：（中央电视台2001年2月27日报道）

从2000年7月1日开始，北京、上海、广州三大城市率先实施《车用无铅汽油》的环保标准，日前国家质量技术监督局在三地进行了新汽油标准实施后的第一次国家监督抽查。

这次共抽查了三市的32家加油站的车用汽油产品，抽样合格率为75%，其中90号无铅车用汽油抽查了15批次，合格10批次，93号车用无铅汽油抽查了17批次，合格14批次。……

本次抽查发现的主要问题是烯烃含量超标，新标准规定，烯烃含量不大于35%，不合格的样品中有5个批次产品的烯烃含量都大于40%，而且主要出在90号汽油上。

报道： （中新社北京2002年12月10日报道）

从上世纪八十年代以来，中国7至17岁青少年男女平均身高分别增长了6.9和5.5厘米，体重分别增长了6.6和4.5公斤。有关专家称，这显示出中国青少年的体格发育呈显著增长趋势，……儿童、青少年的生长长期趋势，已经超过了发达国家二战后出现的生长加速水平。

统计推断：通过对样本的处理和分析，得出与总体参数相关的结论。

统计推断包括参数估计和假设检验两部分内容。

一、实例及其分析

示例一：吸烟对血压有影响吗？

对吸烟组66人，不吸烟组62人：分别监测和测量：

- 24小时收缩压（24hSBP）和舒张压（24hDBP）
- 白天 (6Am-10Pm) 收缩压（dSBP）和舒张压 (dDBP)
- 夜间 (10Pm-6Am) 收缩压（nSBP）和舒张压 (nDBP)

然后分别计算每类的样本均值和标准差：

	吸烟组均值	吸烟组标准差	不吸烟组均值	不吸烟组标准差
24hSBP(mmHg)	119.35	10.77	114.79	8.28
24hDBP(mmHg)	76.83	8.45	72.87	6.20
dSBP(mmHg)	122.70	11.36	117.60	8.71
dDBP(mmHg)	79.52	8.75	75.44	6.80
nSBP(mmHg)	109.95	10.78	107.10	10.11
nDBP(mmHg)	69.35	8.60	65.84	7.03

问题：

1）任何一个考察的时段，吸烟和不吸烟群体的血压的**真值**分别是多少？ **（参数估计）**

2）吸烟和不吸烟群体的血压的**真值**是否有区别? **（假设检验）**

示例二：如何制定汽油供货合同？

某炼油厂（甲方）向加油站（乙方）成批（车次）供货，双方制定了相关的产品质量监控合同，要求含硫量不超过 0.08%. 若双方商定每批抽检10辆车，10个含硫量数据(%)：

0.0864 0.0744 0.0864 0.0752 0.0760
0.0954 0.0936 0.1016 0.0800 0.0880

1）只根据这些数据推断乙方是否应接受该批汽油；如果甲方是可靠的供货商，并且对产品的稳定性提供了进一步的信息，乙方对应的策略有什么变化? (假设检验，第一类错误)

2）现乙方与一新炼油厂（丙方）谈判，并且风闻丙方有用含硫量 0.086% 的汽油顶替合格品的前科，那么如果乙方沿用与甲方订的合同, 会有什么后果? (假设检验，第二类错误)

二、参数估计

参数估计：利用样本统计量对总体参数进行估计，

分点估计和区间估计两种。

- **点估计——矩法、极大似然法**
- **点估计的评价标准**
- **区间估计——总体均值和总体方差**
- **参数估计的MATLAB实现**

点估计：矩法（moment）

点估计：用样本统计量确定总体参数的一个数值

矩法　对总体均值（数学期望）μ，方差σ²的点估计：

$$\bar{x}=\frac{1}{n}\sum_{i=1}^{n}x_i \qquad A_2=\frac{1}{n}\sum_{i=1}^{n}(x_i-\bar{x})^2 \quad \text{即} s_1{}^2$$

$$\hat{\mu}=\bar{x},\ \hat{\sigma}^2=A_2$$

密度 / 分布

一般定义

设总体X具有已知的概率函数$p(x,\theta_1,\ldots,\theta_k)$，$x_1,x_2,\ldots,x_n$是样本，假定总体的$k$阶原点矩$\mu_k$存在，若 $\theta_j=\theta_j(\mu_1,\ldots,\mu_k)$，则可给出诸$\theta_j$的矩法估计为

$$\hat{\theta}_j=\theta_j(a_1,\cdots,a_k),\quad j=1,\cdots,k,$$

其中$a_j=\frac{1}{n}\sum_{i=1}^{n}x_i^j$为$j$阶样本原点矩.

$$\boxed{\mu_j\equiv E(X^j)}$$

点估计：矩法

例1 (上页)对总体X的均值$\theta_1=\mu$，方差$\theta_2=\sigma^2$的点估计：

一阶原点矩 μ_1 = EX = μ ➔ $\mu = \mu_1 = \theta_1$

二阶原点矩 μ_2 = E(X²)= E(X-EX)² +(EX)² = $\sigma^2 + \mu^2$

➔ $\theta_2 = \sigma^2 = \mu_2 - \mu_1^2$

$$\hat{\mu} = \frac{1}{n}\sum_{i=1}^{n} X_i = \bar{X}$$ $$\hat{\sigma}^2 = \frac{1}{n}\sum_{i=1}^{n} X_i^2 - \bar{X}^2 = s_1^2 = \frac{1}{n}\sum_{i=1}^{n}(X_i - \bar{X})^2$$

例2 对指数分布总体X：EX =λ, DX = Var(X) =λ²

$$\hat{\lambda} = \bar{x}$$ $$\hat{\lambda} = s_1$$

矩法的缺陷：

不唯一（常取低阶矩）

点估计：极大似然估计

(MLE：Maximum Likelihood Estimation)

极大似然法 给定的样本 $(x_1, x_2, \cdots, x_n)$ 和总体的概率密度函数 $p(x,\theta)$，求 $\hat{\theta}$ 满足

$$L(\widehat{\theta}) = \max_{\theta} L(x_1, x_2, \cdots, x_n; \theta) = \max_{\theta} \prod_{i=1}^{n} p(x_i; \theta)$$

(L称为**似然函数**) **也不一定唯一**

例1 N(μ，σ²)：未知参数 μ，σ² 的MLE：同矩法(证略)

例2 U(a，b)： 未知参数 a，b 的MLE：与矩法不同

$$L(a,b) = \prod_{i=1}^{n} \frac{1}{b-a} = \frac{1}{(b-a)^n} \Rightarrow \begin{array}{l} \hat{a}_L = x_{(1)} = \min(X_1, X_2, \cdots, X_n) \\ \hat{b}_L = x_{(n)} = \max(X_1, X_2, \cdots, X_n) \end{array}$$

单选题 2分

设置

设随机变量 X 的概率密度为：$f(x)=\begin{cases}\dfrac{2x}{\theta^2}, & 0<x<\theta, \\ 0, & \text{其他}\end{cases}$

其中未知参数 $\theta>0$，$X_1,\cdots,X_n$ 是来自 X 的样本，

则 θ 的矩估计和极大似然估计分别是

A $\frac{2}{3}\overline{X}$ $\min(X_1,X_2,\cdots,X_n)$

B $\frac{2}{3}\overline{X}$ $\max\{X_1,X_2,\cdots,X_n\}$

C $\frac{3}{2}\overline{X}$ $\min(X_1,X_2,\cdots,X_n)$

D $\frac{3}{2}\overline{X}$ $\max\{X_1,X_2,\cdots,X_n\}$

E 以上都不对

提交

评价估计优劣的标准

待估参数理想值θ，参数估计随机变量 $\hat{\theta}$

无偏性(unbiased) $E\hat{\theta} = \theta$

对于正态总体 $N(\mu,\sigma^2)$ 的样本（对于一般总体也对）

$$E\bar{x} = \frac{1}{n}\sum_{i=1}^{n} Ex_i = \frac{n\mu}{n} = \mu \qquad EA_2 = \frac{n-1}{n}\sigma^2$$

均值的参数估计无偏，方差参数估计有偏

改进 $s^2 = \dfrac{1}{n-1}\sum_{i=1}^{n}(x_i - \bar{x})^2$

对方差进行参数估计无偏

$A_2 = s_1^2$：渐进无偏

有效性 在所有无偏估计量中

$$D\hat{\theta} = E(\hat{\theta}-\theta)^2 \qquad D\theta^* = \min_{\hat{\theta}} E(\hat{\theta}-\theta)^2$$

θ^* 称为θ的有效估计量（最小方差无偏估计）

正态总体：$\bar{x}$ 和 s^2 分别是 μ 和 σ^2 的最小方差估计

非正态总体：$\bar{x}$ 和 s^2 一般不是 μ 和 σ^2 的最小方差估计

一致性 (consistent, **常称相容性、(弱)相合性**)

如果对任给的 $\varepsilon > 0$，满足 $\lim_{n\to\infty} P(|\hat{\theta}_n-\theta|<\varepsilon)=1$

$\hat{\theta}_n$ 称为θ的一致估计量（依概率收敛）

样本均值 $\bar{x}$ 和方差 s^2 分别是 μ 和 σ^2 的一致无偏估计

多选题 2分

设置

（简单随机样本）

设 ξ_1，ξ_2 是取自正态母体 $N(\mu,1)$ 的一个容量为 2 的子样，

$\hat{\mu}_1 = \frac{2}{3}\xi_1 + \frac{1}{3}\xi_2$，$\hat{\mu}_2 = \frac{1}{4}\xi_1 + \frac{3}{4}\xi_2$，则作为母体均值的点估计，

A $\hat{\mu}_1$，$\hat{\mu}_2$ 均是 μ 的无偏估计

B $\hat{\mu}_1$ 比 $\hat{\mu}_2$ 更有效

C $\hat{\mu}_2$ 比 $\hat{\mu}_1$ 更有效

D $D\hat{\mu}_1 > D\hat{\mu}_2$ （D 表示方差）

E 以上都不对

提交

区间估计

总体的待估参数θ，估计量 $\hat{\theta}$，求区间 $[\hat{\theta}_1,\hat{\theta}_2]$，使θ满足

$$P(\hat{\theta}_1 \le \theta \le \hat{\theta}_2) = 1-\alpha \quad (0<\alpha<1, \text{给定})$$

$[\hat{\theta}_1,\hat{\theta}_2]$：θ 的置信区间　　$\hat{\theta}_1,\hat{\theta}_2$：置信下限和置信上限

$1-\alpha$：置信概率或置信水平　　α：**显著性水平**

$\alpha = 0.05$ ⇨ 由样本得到的置信区间以 0.95 的概率包含了待估参数θ

置信区间越小，估计精度越高 ⟷ **二者矛盾** ⟷ **置信水平越大，可信程度越高**

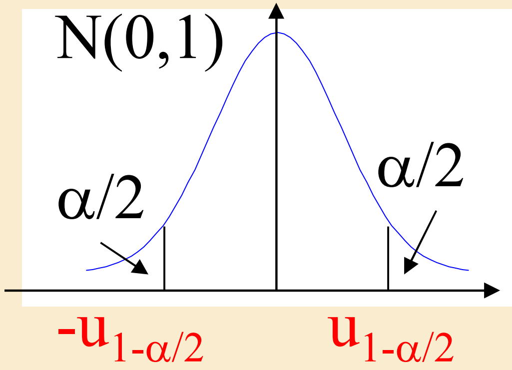
在一定置信水平下使置信区间尽量小

总体均值μ的区间估计

假设总体服从正态分布N(μ , σ²)

1) 总体方差σ^2已知，估计均值μ，

置信水平 $1-\alpha$



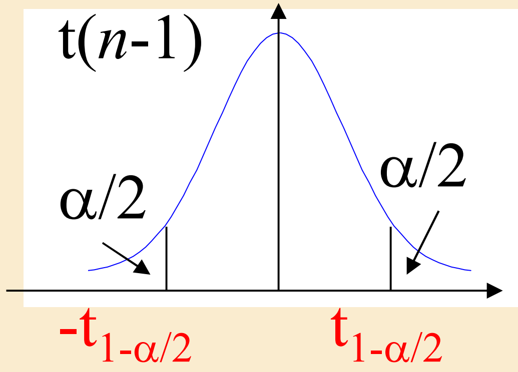
样本均值 $\bar{x}$ $\quad z=\dfrac{\bar{x}-\mu}{\sigma/\sqrt{n}} \sim N(0,1)$

$N(0,1)$的$1-\alpha/2$分位数$u_{1-\alpha/2}$满足 $P(|z| \le u_{1-\alpha/2}) = 1-\alpha$

$$\Rightarrow P(\bar{x}-u_{1-\alpha/2}\frac{\sigma}{\sqrt{n}} \le \mu \le \bar{x}+u_{1-\alpha/2}\frac{\sigma}{\sqrt{n}}) = 1-\alpha$$

置信区间 $\quad [\bar{x}-u_{1-\alpha/2}\dfrac{\sigma}{\sqrt{n}},\ \bar{x}+u_{1-\alpha/2}\dfrac{\sigma}{\sqrt{n}}]$

2) 总体方差σ^2未知，估计均值μ，置信水平 $1-\alpha$



样本均值 $\bar{x}$

样本均方差s

$$\frac{\bar{x}-\mu}{s/\sqrt{n}} \sim t(n-1)$$

$t(n-1)$的$1-\alpha/2$分位数$t_{1-\alpha/2}$满足：

$$P(\bar{x}-t_{1-\alpha/2}\frac{s}{\sqrt{n}} \le \mu \le \bar{x}+t_{1-\alpha/2}\frac{s}{\sqrt{n}}) = 1-\alpha$$

置信区间 $[\bar{x}-t_{1-\alpha/2}\frac{s}{\sqrt{n}},\ \bar{x}+t_{1-\alpha/2}\frac{s}{\sqrt{n}}]$

与σ^2已知比较 $[\bar{x}-u_{1-\alpha/2}\frac{\sigma}{\sqrt{n}},\ \bar{x}+u_{1-\alpha/2}\frac{\sigma}{\sqrt{n}}]$

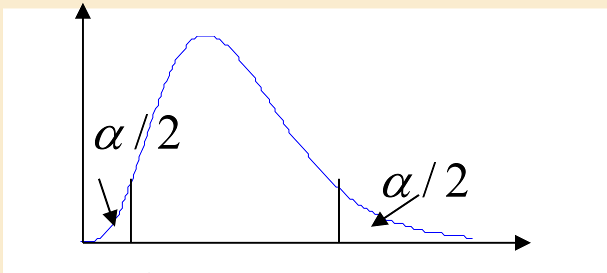
$$\alpha = 0.05,\ u_{1-\alpha/2} = 1.96,\ \ t_{1-\alpha/2} = 2.06\,(n=25)$$

总体均值方差σ²的区间估计

假设总体服从正态分布N(μ , σ²)

样本方差s^2

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



$\chi^2_{\alpha/2}$ $\chi^2_{1-\alpha/2}$

χ^2分布的$\alpha/2$分位数 $\chi^2_{\alpha/2}$和 $1-\alpha/2$分位数 $\chi^2_{1-\alpha/2}$满足

$$P(\chi^2_{\alpha/2} \le \frac{(n-1)s^2}{\sigma^2} \le \chi^2_{1-\alpha/2}) = 1-\alpha$$

σ^2的置信区间 $$[\frac{(n-1)s^2}{\chi^2_{1-\alpha/2}}, \frac{(n-1)s^2}{\chi^2_{\alpha/2}}]$$

小结：总体均值μ，方差σ²的区间估计

μ的置信区间(σ²已知) $[\bar{x}-u_{1-\alpha/2}\dfrac{\sigma}{\sqrt{n}},\ \bar{x}+u_{1-\alpha/2}\dfrac{\sigma}{\sqrt{n}}]$

μ的置信区间(σ²未知) $[\bar{x}-t_{1-\alpha/2}\dfrac{s}{\sqrt{n}},\ \bar{x}+t_{1-\alpha/2}\dfrac{s}{\sqrt{n}}]$

σ²的置信区间 $[\dfrac{(n-1)s^2}{\chi^2_{1-\alpha/2}},\ \dfrac{(n-1)s^2}{\chi^2_{\alpha/2}}]$

μ的置信区间长度 $L_\mu = 2t_{1-\frac{\alpha}{2}}s/\sqrt{n} \approx 2u_{1-\frac{\alpha}{2}}s/\sqrt{n}$ (n大)

σ²的置信区间长度 $L_{\sigma^2} = (n-1)s^2(1/\chi^2_{\alpha/2} - 1/\chi^2_{1-\alpha/2})$

给定α：n越大，通常L_μ越小，估计精度越高；L_{σ^2} 呢?

参数估计的MATLAB实现

```
[mu sigma muci sigmaci] = normfit(x,alpha)
```

其中: x为样本

alpha为显著性水平（缺省时为0.05)

返回值：mu-- 均值μ的点估计

sigma---标准差σ的点估计

muci--均值μ的区间估计

sigmaci---标准差σ的区间估计

示例-----学生身高数据处理

50名17岁城市男性学生身高（单位：cm）：

170.1 179.0 171.5 173.1 174.1 177.2 170.3 176.2 163.7 175.4
163.3 179.0 176.5 178.4 165.1 179.4 176.3 179.0 173.9 173.7
173.2 172.3 169.3 172.8 176.4 163.7 177.0 165.9 166.6 167.4
174.0 174.3 184.5 171.9 181.4 164.6 176.4 172.4 180.3 160.5
166.2 173.5 171.7 167.9 168.7 175.6 179.6 171.6 168.1 172.2

计算结果（α=0.05 ）

	身高
均值点估计	172.7040
均值区间估计	（171.1777, 174.2303）
标准差点估计	5.3707
标准差区间估计	（4.4863，6.6926）

Demo1201.m

投票 最多可选5项

设置

根据某地区关于工资的样本资料，估计出的该地区平均工资的95%的置信区间为[700，1500]，则下列说法准确的有

A 该地区平均工资有95%的可能性落到该置信区间

B 该置信区间的误差不会超过5%

C 该置信区间有95%的概率包含该地区的平均工资

D 该置信区间要么包含该地区的平均工资，要么不包含该地区的平均工资

E 在用同样方法构造该地区平均工资的多个置信区间中，包含平均工资的区间比率为95%

提交

三、假设检验

- 总体均值的假设检验
- 总体方差的假设检验
- 两总体的假设检验
- 0-1分布总体均值的假设检验
- 总体分布正态性检验
- 假设检验的MATLAB实现

例　**甲方产品：$x \sim N(50,1)$，批量供给乙方。**

对于每批产品μ=50是否成立，双方商定检验方案。

- 每批抽取25件测量，计算均值 $\bar{x}$
- 制订数量标准δ，若 $|\bar{x}-50| \le \delta$，认为

μ=50成立，接受该批产品（合格品）；否则，拒绝。

- 商定水平α，使合格品被错误拒绝的概率

不超过α　(通常α =0.05)。

解 $z = \dfrac{\bar{x}-\mu}{\sigma/\sqrt{n}} \sim N(0,1)$　$P(|z| \le 2) = 0.95$　$P(|\bar{x}-\mu| \le 2\sigma/\sqrt{n}) = 0.95$

⇨ $\delta = 2\sigma/\sqrt{n} = 0.4$　　**当** $|\bar{x}-50| \le 0.4$ **接受；否则，拒绝。**

总体均值的假设检验

已有样本(容量n，均值 $\bar{x}$,标准差s),要对总体均值μ是否等于给定值μ_0进行检验(假定总体服从正态分布)

假设 $H_0: \mu = \mu_0$; $H_1: \mu \neq \mu_0$

称H_0为**原假设（或零假设）**，H_1为**备选假设**，

二择一：接受**(不拒绝)**H_0；拒绝H_0，即接受**(不拒绝)**H_1

显著性水平$\alpha \sim H_0$成立时被错误拒绝的概率。

总体方差σ^2已知

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

$$P(|z| \le u_{1-\alpha/2}) = 1-\alpha$$

⇨ $|z| \le u_{1-\alpha/2}$时接受H_0;

否则拒绝H_0(接受H_1).

称z**检验**或u**检验**　**接受域**　**拒绝域(否定域、临界域)**

总体方差 σ^2未知

$$t=\frac{\bar{x}-\mu_0}{s/\sqrt{n}}\sim t(n-1) \qquad P(|t|\le t_{1-\alpha/2})=1-\alpha$$

$|t|\le t_{1-\alpha/2}$时接受H_0;否则拒绝H_0(接受H_1).

称t检验

常用: α =0.05→$u_{1\text{-}\alpha/2}$=1.96; α =0.01→$u_{1\text{-}\alpha/2}$=2.575

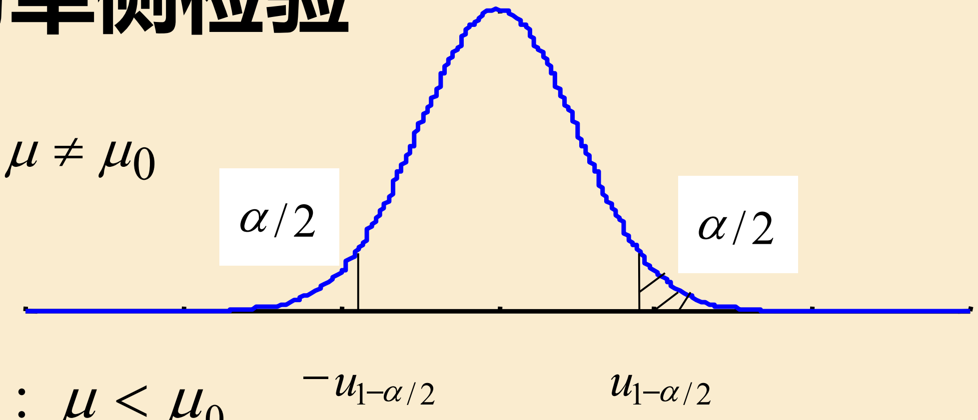
当n较大时(n>30) $t_{1\text{-}\alpha/2}$ 与$u_{1\text{-}\alpha/2}$相近.

思考　设从一个样本得到 z=2.2，那么若取α=0.05，将拒绝 H_0；若取α=0.01，将接受 H_0。你怎样评价这两个不同的结果？

α是错误地拒绝 H_0 的概率，α不是越小越好吗?

总体均值的假设检验

双侧检验与单侧检验



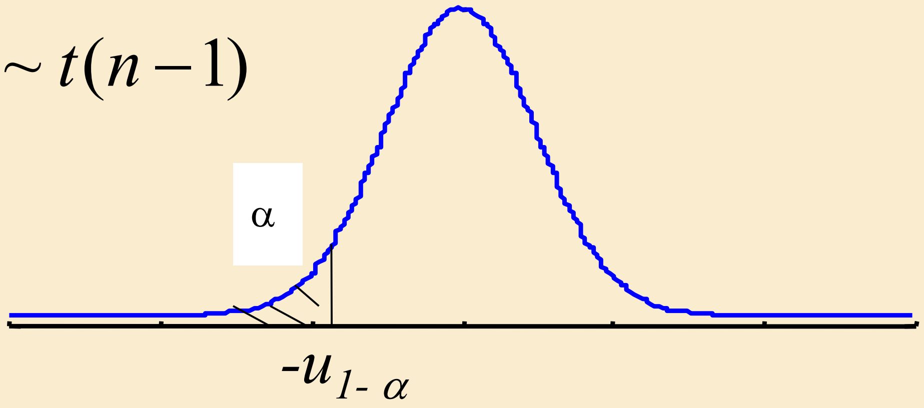
双侧检验 $H_0: \mu = \mu_0; \quad H_1: \mu \neq \mu_0$

单侧检验 $H_0: \mu \geq \mu_0; \quad H_1: \mu < \mu_0$

$$z = \frac{\bar{x} - \mu_0}{\sigma/\sqrt{n}} \sim N(0,1) \qquad t = \frac{\bar{x} - \mu_0}{s/\sqrt{n}} \sim t(n-1)$$

$z \geq u_\alpha (= -u_{1-\alpha})$时接受 H_0;

否则拒绝 H_0(接受 H_1).



单侧检验 $H_0: \mu \leq \mu_0; \quad H_1: \mu > \mu_0$

总体均值假设检验小结

总体方差σ^2已知

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

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

⇨ $|z| \leq u_{1-\alpha/2}$时接受H_0;否则拒绝H_0.

$H_0: \mu \geq \mu_0$; $H_1: \mu < \mu_0$ $z \geq u_\alpha (= -u_{1-\alpha})$时接受$H_0$;否则拒绝$H_0$.

总体方差σ^2未知

$H_0: \mu = \mu_0$; $H_1: \mu \neq \mu_0$ $|t| \leq t_{1-\alpha/2}$时接受H_0;否则拒绝H_0.

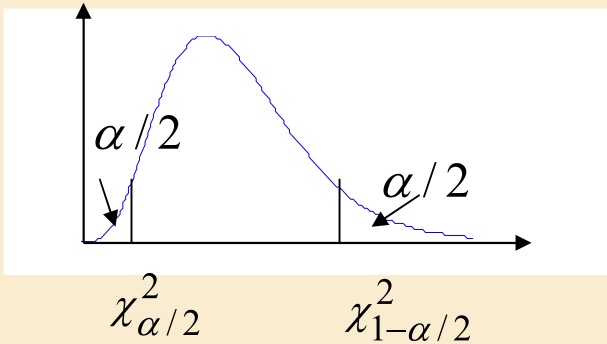
$$t = \frac{\bar{x} - \mu_0}{s / \sqrt{n}} \sim t(n-1)$$ ⇨

$H_0: \mu \geq \mu_0$; $H_1: \mu < \mu_0$ $z \geq t_\alpha (= -t_{1-\alpha})$时接受$H_0$;否则拒绝$H_0$.

总体方差的假设检验

双侧检验 $H_0: \sigma^2=\sigma_0^2,\quad H_1: \sigma^2\neq\sigma_0^2$

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



$\chi^2_{\alpha/2}\le\chi^2\le\chi^2_{1-\alpha/2}$ 时接受H_0;否则拒绝H_0.

单侧检验 $H_0: \sigma^2\ge\sigma_0^2,\quad H_1: \sigma^2<\sigma_0^2$

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

问题

粮食加工厂的一台自动装包机的设定值为：每包装50公斤，标准差0.3。现抽查了20包，得到如下数据

49.8	50.1	50.5	49.7	49.0	50.0	50.3	50.0	49.9	49.9
50.5	49.2	49.7	49.8	50.1	50.0	50.3	50.2	50.4	50.1

问该包装机运转是否正常（这里指是否稳定）？

单侧检验 $H_0: \sigma^2 \le 0.09,\quad H_1: \sigma^2 > 0.09$ 取α=0. 05

计算样本方差 $s^2=0.1483$

$$\chi^2 = \frac{(n-1)s^2}{\sigma_0^2} = 31.3056 \quad > \chi^2_{1-\alpha} = 30.1435$$

拒绝H_0

取α=0. 02 试试看！

两总体的假设检验——均值

总体方差σ_1^2, σ_2^2已知 $H_0: \mu_1 = \mu_2$, $H_1: \mu_1 \neq \mu_2$

$$X \sim N(\mu_1, \sigma_1^2) \qquad Y \sim N(\mu_2, \sigma_2^2) \quad \Rightarrow \quad \frac{(\bar{x} - \mu_1) - (\bar{y} - \mu_2)}{\sqrt{\frac{\sigma_1^2}{n_1} + \frac{\sigma_2^2}{n_2}}} \sim N(0,1)$$

若H_0成立 $\Rightarrow$ $$z = \frac{\bar{x} - \bar{y}}{\sqrt{\frac{\sigma_1^2}{n_1} + \frac{\sigma_2^2}{n_2}}} \sim N(0,1)$$

给定显著性水平α

取$N(0,1)$的$1-\alpha/2$分位数$u_{1-\alpha/2}$ $\Rightarrow$ $|z| \le u_{1-\alpha/2}$时接受H_0；否则拒绝H_0（接受H_1）.

总体方差σ_1^2, σ_2^2未知，但可假定$\sigma_1^2=\sigma_2^2$

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

$$\frac{(\bar{x}-\mu_1)-(\bar{y}-\mu_2)}{\sqrt{\dfrac{s^2}{n_1}+\dfrac{s^2}{n_2}}}\sim t(n_1+n_2-2),\quad s^2=\frac{(n_1-1)s_1^2+(n_2-1)s_2^2}{n_1+n_2-2}$$

H_0成立 $\Rightarrow$ $$t=\frac{\bar{x}-\bar{y}}{\sqrt{\dfrac{s^2}{n_1}+\dfrac{s^2}{n_2}}}\sim t(n_1+n_2-2)$$

取分位数 $t_{1-\alpha/2}$, $|t|\le t_{1-\alpha/2}$ 时接受 H_0;否则拒绝 H_0(接受 H_1).

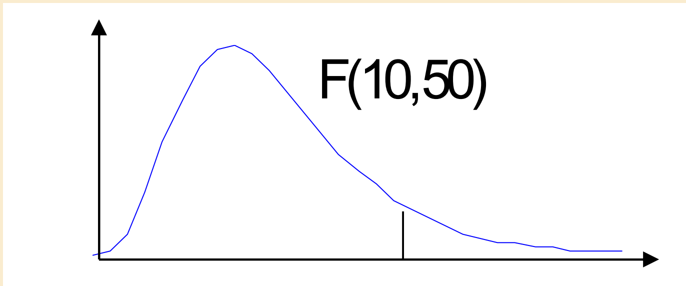
两总体的假设检验——方差

$X \sim N(\mu_1, \sigma_1^2)$ $Y \sim N(\mu_2, \sigma_2^2)$

已知2个样本n_1, n_2, s_1^2, s_2^2 $\Rightarrow$ $\dfrac{s_1^2/\sigma_1^2}{s_2^2/\sigma_2^2} \sim F(n_1-1, n_2-1)$

双侧假设检验 $H_0: \sigma_1^2 = \sigma_2^2, \quad H_1: \sigma_1^2 \neq \sigma_2^2$

若H_0成立 $\dfrac{s_1^2}{s_2^2} \sim F(n_1-1, n_2-1)$



$F_{1-\alpha/2}$

取 $s_1^2 \geq s_2^2$，记 $F = \dfrac{s_1^2}{s_2^2}$

给定α，取$F(n_1\text{-}1, n_2\text{-}1)$的1- α/2分位数$F_{1-\alpha/2}$

当 $F \leq F_{1-\alpha/2}$ 时接受 H_0；
否则拒绝 H_0（接受 H_1）

置信水平仍为1- α

0-1分布总体均值的假设检验

问题

甲方向乙方成批供货，双方商定废品率不超过 3%。今从一批中抽取 100 件，发现有 5 件废品，问乙方是否应接受这批产品。（设$\alpha=0.05$）

分析

总体服从0-1分布：X=0—合格品; X=1—废品

X的均值$\mu=p$ (废品率)，X的方差$\sigma^2=p(1-p)$

样本容量n，均值 $\bar{x}$ (平均废品率)

n充分大时近似地有 $z=\dfrac{\bar{x}-\mu}{\sigma/\sqrt{n}}\sim N(0,1)$

总体废品率的

假设检验(双侧) $H_0: p=p_0,\quad H_1: p\neq p_0$

H_0成立时 $$z=\frac{\bar{x}-\mu}{\sigma/\sqrt{n}}=\frac{\bar{x}-p_0}{\sqrt{\frac{p_0(1-p_0)}{n}}}\sim N(0,1)$$

取$N(0,1)$的1-$\alpha/2$分位数$u_{1\text{-}\alpha/2}$ $\quad P(|z|\le u_{1-\alpha/2})=1-\alpha$

$|z|\le u_{1-\alpha/2}$ 时接受 H_0;否则拒绝 H_0(接受 H_1).

假设检验(单侧) $H_0: p\le p_0,\quad H_1: p>p_0$

$z\le u_{1-\alpha}$ 时接受 H_0;否则拒绝 H_0(接受 H_1).

$\bar{x}=5/100,\ p_0=0.03,\ n=100,\ z=1.17<u_{1-\alpha}=u_{0.95}=1.65$

问题

H_0成立，乙方应接受那批产品.

问：如果将α提高，会有什么结果?

总体分布的正态性检验

1） Jarque-Bera检验

正态分布的**偏度**$g_1=0$，**峰度**$g_2=3$，对于一个样本计算其g_1和g_2，若样本来自正态总体，则g_1和g_2应分别在0和3附近。基于这个思想，构造一个包含g_1，g_2的 χ^2 统计量。

2） Kolmogorov-Smirnov检验

通过样本的经验分布函数与给定分布函数的比较，推断该样本是否来自给定分布函数的总体。该正态性检验只能做标准正态检验 。

3） Lilliefors检验

它将Kolmogorov-Smirnov检验改进用于一般的正态性检验。

假设检验的MATLAB实现

假设检验		统计量	检验规则	MATLAB 命令
单个总体均值（σ^2 已知）	$H_0: \mu=\mu_0$ $H_1: \mu\neq\mu_0$	$z=\dfrac{\bar{x}-\mu_0}{\sigma/\sqrt{n}}\sim N(0,1)$	$\lvert z\rvert \le u_{1-\alpha/2}$ 接受 H_0 （z 检验）	h=ztest(x,mu,sigma) [h,sig,ci,zval]=ztest(x,mu,sigma,alpha,tail)
单个总体均值（σ^2 未知）	$H_0: \mu=\mu_0$ $H_1: \mu\neq\mu_0$	$t=\dfrac{\bar{x}-\mu_0}{s/\sqrt{n}}\sim t(n-1)$	$\lvert t\rvert \le t_{1-\alpha/2}$ 接受 H_0 （t 检验）	h=ttest(x,mu) [h,sig,ci]=ttest(x,mu,alpha,tail)
单个总体方差	$H_0: \sigma^2=\sigma_0^2$ $H_1: \sigma^2\neq\sigma_0^2$	$\chi^2=\dfrac{(n-1)s^2}{\sigma_0^2}\sim\chi^2(n-1)$	$\chi^2_{\alpha/2}\le\chi^2\le\chi^2_{1-\alpha/2}$ 接受 H_0	无

假设检验		统计量	检验规则	MATLAB 命令
两个总体均值（σ_1^2,σ_2^2 已知）	$H_0: \mu_1=\mu_2$ $H_1: \mu_1\neq\mu_2$	$z=\dfrac{\bar{x}-\bar{y}}{\sqrt{\dfrac{\sigma_1^2}{n_1}+\dfrac{\sigma_2^2}{n_2}}}\sim N(0,1)$	$\lvert z\rvert\le u_{1-\alpha/2}$ 接受 H_0	无 **也可用ttest2**
两个总体均值（$\sigma_1^2=\sigma_2^2$ 未知）	$H_0: \mu_1=\mu_2$ $H_1: \mu_1\neq\mu_2$	$t=\dfrac{\bar{x}-\bar{y}}{\sqrt{\dfrac{s^2}{n_1}+\dfrac{s^2}{n_2}}}\sim t(n_1+n_2-2)$ $s^2=\dfrac{(n_1-1)s_1^2+(n_2-1)s_2^2}{n_1+n_2-2}$	$\lvert t\rvert\le t_{1-\alpha/2}$ 接受 H_0	h=ttest2(x,y) [h,sig,ci]=ttest2(x,y,alpha,tail)
两个总体方差	$H_0: \sigma_1^2=\sigma_2^2$ $H_1: \sigma_1^2\neq\sigma_2^2$	$F=\dfrac{s_1^2}{s_2^2}\sim F(n_1-1,n_2-1),\quad s_1^2\ge s_2^2$	$F\le F_{1-\alpha/2}$ 接受 H_0	无

新的命令格式：ttest2(x, y, ’Alpha’, 0.05, ‘Tail’, ‘Left’, 'Vartype','unequal')

假设检验		统计量	检验规则	MATLAB 命令
0-1 分布总体均值	$H_0:\ p = p_0$ $H_1:\ p \neq p_0$	$z = \dfrac{\bar{x} - p_0}{\sqrt{p_0(1-p_0)/n}}$	$\lvert z \rvert \le u_{1-\alpha/2}$ 接受 H_0	无
总体分布正态性	H_0：总体服从 $N(\mu,\sigma^2)$	略	略	h =jbtest(x)
				[h,p,jbstat,cv] =jbtest(x,alpha)
	H_0：总体服从 $N(0,1)$	略	略	h =kstest(x)
	H_0：总体服从 $N(\mu,\sigma^2)$	略	略	h =lillietest(x)
				[h,p,lstat,cv]= lillietest(x,alpha)

MATLAB命令使用说明

输入参数 x 是样本（n 维数组），mu 是 H_0 中的 μ_0，sigma 是总体标准差 σ，alpha 是显著性水平 α（缺省时设定为 0.05），tail 是对双侧检验和两个单侧检验的标识，用备选假设 H_1 确定：H_1 为 $\mu \neq \mu_0$ 时令 tail=0（可缺省）；H_1 为 $\mu > \mu_0$ 时令 tail=1；H_1 为 $\mu < \mu_0$ 时令 tail=-1。

输出参数h=0表示接受H_0, h=1表示拒绝H_0,

sig (P值)标示对假设的接受和拒绝程度。

ci给出置信区间，zval是样本统计量z的值。

用N(5,1)随机数产生n=100的样本，在总体方差未知的情况下分别取α=0.05和α=0.01检验总体均值$\mu \geq 5.2$。

$$H_0:\ \mu \geq 5.2,\quad H_1:\ \mu < 5.2$$

```
x = normrnd(5,1,100,1);
m = mean(x),
[h1,sig1,ci1] = ttest(x,5.2,0.05,-1)
[h2,sig2,ci2] = ttest(x,5.2,0.01,-1)
```

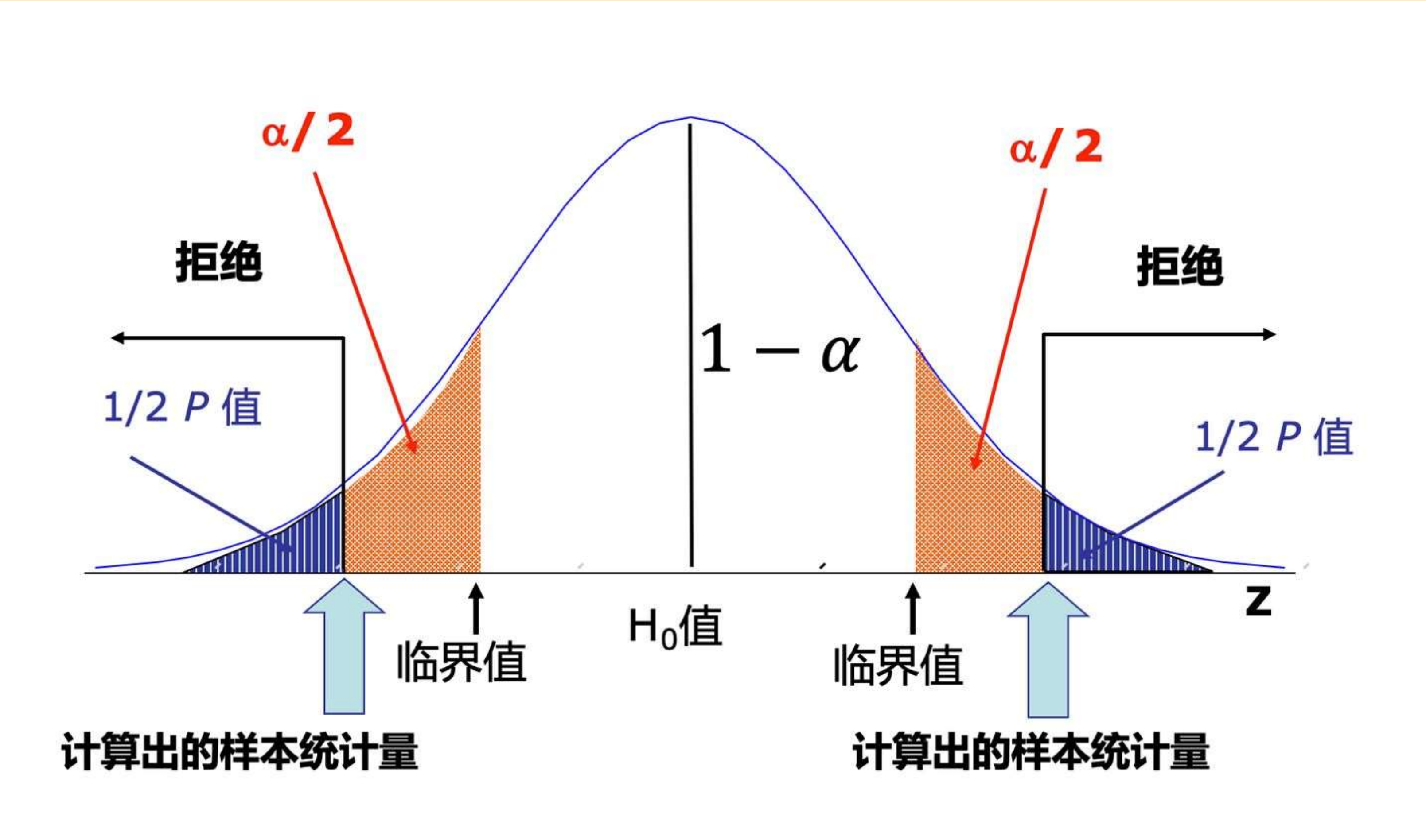
demo1202.m

计算结果

```
m = 5.0111,
[h1,sig1,ci1] =   1   0.0343     -Inf   5.1815
[h2,sig2,ci2] =   0   0.0343     -Inf   5.2537
```

可知在α=0.05下拒绝H_0（此时sig1<α），μ的区间估计（-∞ 5.1815]不包含5.2；而在α=0.01下接受H_0（此时sig2>α），μ的区间估计（-∞ 5.2537]包含5.2。

P值的含义：双侧检验为例



单选题 1分

设置

假设检验中的P值反映的是:

A 拒绝域的大小

B 事先给定的显著性水平的大小

C 拒绝原假设的对或错

D 如果原假设 H_0 为真，所得到的样本结果会像实际观测结果那么极端或更极端的概率

E 以上都不对

提交

四. 实例的求解

示例：吸烟对血压的影响

测量了吸烟组（66人）和不吸烟组（62人）两组人群的6项血压指标

	吸烟组均值	吸烟组标准差	不吸烟组均值	不吸烟组标准差
24hSBP(mmHg)	119.35	10.77	114.79	8.28
24hDBP(mmHg)	76.83	8.45	72.87	6.20
dSBP(mmHg)	122.70	11.36	117.60	8.71
dDBP(mmHg)	79.52	8.75	75.44	6.80
nSBP(mmHg)	109.95	10.78	107.10	10.11
nDBP(mmHg)	69.35	8.60	65.84	7.03

两个总体均值的假设检验

$$H_0: \mu_1 = \mu_2; \quad H_1: \mu_1 \neq \mu_2$$

```
function
[h,sig]=pttest2(xbar,ybar,s1,s2,m,n,alpha,tail)
spower=((m-1)*s1^2+(n-1)*s2^2)/(m+n-2);
t=(xbar-ybar)/sqrt(spower/m+ spower/n);
if tail==0
    a=tinv(1-alpha/2,m+n-2);
    sig =2*(1-tcdf(abs(t),m+n-2));
    if abs(t)<=a
        h=0;
    else
        h=1;
    end
end
if tail==1
    a=tinv(1-alpha,m+n-2);
    sig =1-tcdf(t,m+n-2);
    if t<=a
        h=0;
    else
        h=1;
    end
end
if tail==-1
    a=tinv(alpha,m+n-2);
   sig =tcdf(t,m+n-2);
    if t>=a
        h=0;
    else
        h=1;
    end
end
```

计算结果（α=0.05）

Demo1203.m

	h	Sig	接受或拒绝H_0
24hSBP(mmHg)	1	8.5097e-003	拒绝
24hDBP(mmHg)	1	3.1860e-003	拒绝
dSBP(mmHg)	1	5.3064e-003	拒绝
dDBP(mmHg)	1	3.9950e-003	拒绝
nSBP(mmHg)	0	1.2597e-001	接受
nDBP(mmHg)	1	1.3027e-002	拒绝

除夜间（10Pm-6Am）平均收缩压（nSBP）外，其余5项指标都拒绝了H_0，于是综合起来可以认为，吸烟对血压的影响显著。

示例：加油站合同制定问题

加油站（乙方）以含硫量不超过0.08%的标准决定是否接受炼油厂（甲方）提供的一批汽油。双方商定每批抽检10辆车，现得到了一批10个含硫量数据

0.0864 0.0744 0.0864 0.0752 0.0760
0.0954 0.0936 0.1016 0.0800 0.0880 (%)

解： 假设检验 $H_0: \mu \le \mu_0 = 0.08; \quad H_1: \mu > \mu_0 = 0.08$

1）显著性水平α=0.05，方差未知的情况下

```
x=[0.0864 …… 0.0880];        % 10个含硫量数据
xbar=mean(x)；               % 样本均值
[h,sig]=ttest(x,0.08,0.05,1)  % t（单侧）检验（设α=0.05）
```

Demo1204.m

计算结果：[h,sig]= 1 0.0424， **拒绝**假设

2）显著性水平α=0.05，标准差为σ=0.01

[h, sig]=**ztest**(x,0.08,0.01,0.05,1)　% z（单侧）检验

计算结果：[h, sig]= 1　0.0357

拒绝假设，乙方不接受该批汽油。

3）显著性水平α=0.05，标准差为σ=0.015

[h, sig]=**ztest**(x,0.08,**0.015**,0.05,1)　% z（单侧）检验

计算结果：[h, sig]= 0　0.1147

乙方**接受**该批汽油。

分析：由于数据量较小（10个），随机性较大。

- 如果生产稳定（标准差为σ=0.01），数据**不能接受**
- 如果生产稳定性略差，这样的数据**可以接受**

4） 若对甲方产品的信任度很高，不妨将显著性水平由α=0.05改为α=0.01， 重新计算。

```
[h,sig]=ttest(x,0.08,0.01,1)
```

计算结果：[h,sig]= 0   0.0424

同一个样本用于同样的假设检验，在不同的显著性水平α下会得到不同的结论。

如何恰当地选取α? 因为α是原本成立的H_0被错误地拒绝的概率，而在假设检验中原假设H_0一般是受保护的，不轻易拒绝它，所以α一般取得很小，并且H_0越可靠，α越小。在上面的问题中甲方一向信誉很好，减小α是合适的。

5） 现乙方与一新炼油厂(丙方)谈判，如沿用与甲方订的合同, 会有什么后果（风闻丙方有用含硫量0.086% 的汽油顶替合格品的前科）。

假设检验 $H_0: \mu \le \mu_0 = 0.08; \quad H_1: \mu > \mu_0 = 0.08$

检验规则 $t = \dfrac{\bar{x} - \mu_0}{s/\sqrt{n}} \le t_{1-\alpha}$ **接受**

$$H_1 \Rightarrow \mu = \mu_2 = 0.086 \qquad t_2 = \frac{\bar{x} - \mu_2}{s/\sqrt{n}} = \frac{\bar{x} - \mu_0 + \mu_0 - \mu_2}{s/\sqrt{n}} = t - \frac{\mu_2 - \mu_0}{s/\sqrt{n}}$$

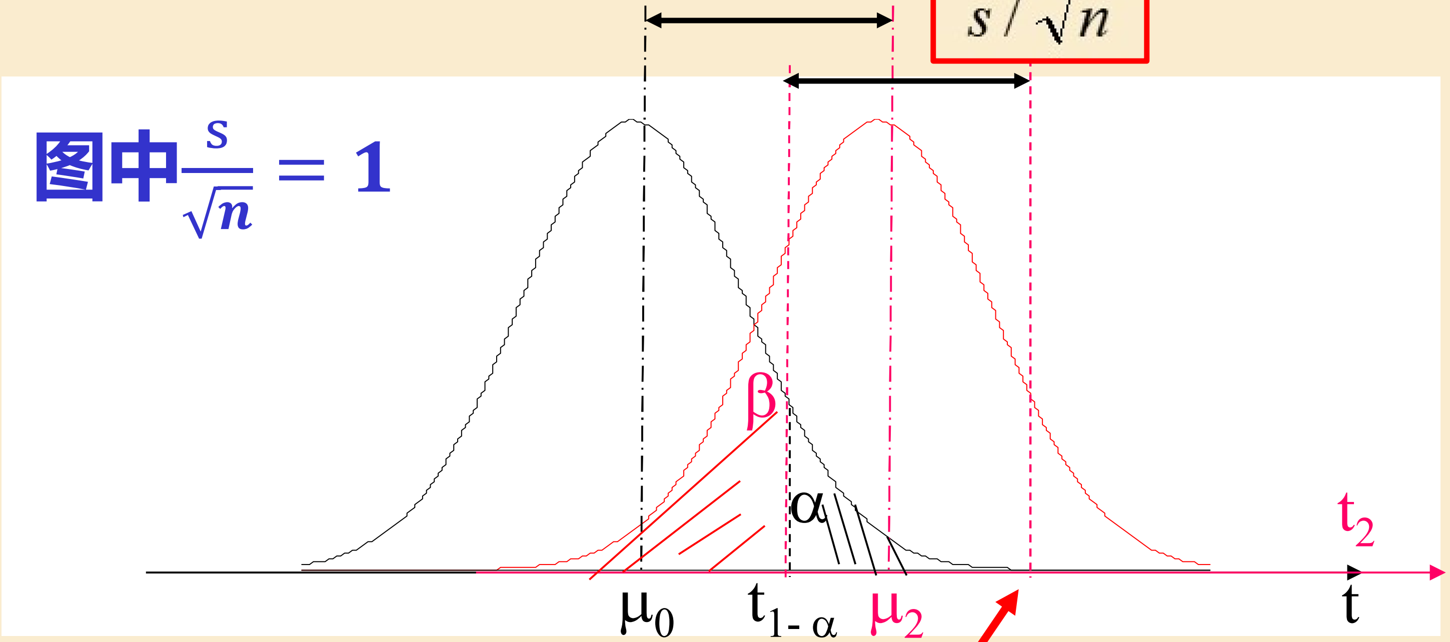
“原本不成立的H_0被接受”（“取伪”)称为**第二类错误**

$$\beta = P(t = \frac{\bar{x} - \mu_0}{s/\sqrt{n}} \le t_{1-\alpha} | H_1)$$

t_β

$$\Rightarrow \quad \beta = P(t_2 < t_{1-\alpha} - \frac{\mu_2 - \mu_0}{s/\sqrt{n}} | H_1) = F_{t(n-1)}(t_{1-\alpha} - \frac{\mu_2 - \mu_0}{s/\sqrt{n}})$$

$$\boxed{\frac{\mu_2-\mu_0}{s/\sqrt{n}}}$$



$$t_\beta = t_{1-\alpha} - \frac{\mu_2-\mu_0}{s/\sqrt{n}}$$

$\alpha\downarrow \rightarrow t_{1-\alpha}\uparrow$, $t_\beta\uparrow \rightarrow \beta\uparrow$

Demo1205.m

显著性水平α	0.01	0.05	0.1	0.2	0.3
第二类错误β	0.7733	0.4211	0.2644	0.1390	0.0846

假设检验中的两类错误

第一类错误“**弃真**”——本应接受的H_0被拒绝, 概率α

第二类错误“**取伪**”——本应拒绝的H_0被接受, 概率β

当样本容量一定时，**二者矛盾**：α减小导致β增加.

通常α选得较小(0.05,0.01), β则较大(具体数值取决于μ_1).

原假设H_0和备选假设H_1是不平等的:

人们保护、偏爱H_0； “歧视” H_1。

实际问题中选择什么样的H_0是重要的

单选题 1分

设置

设$X_1 \ldots X_n$是来自$N(u,1)$的样本，考虑如下的假设检验问题： H_0: $u=2$ ； H_1: $u=3$

若检验由拒绝域 $W=\{\bar{x} \geq 2.6\}$ 确定，

如果要使得检验犯第二类错误的概率小于0.01，则n最小应取多少？

A 34

B 35

C 36

D 37

E 38

提交

按照第二类错误的定义：

$$\beta = P(\bar{x} < 2.6 | H_1) = P\left(\frac{\bar{x}-3}{\sqrt{1/n}} < \frac{2.6-3}{\sqrt{1/n}}\right)$$

$$= \Phi(-0.4\sqrt{n}) = 1 - \Phi(0.4\sqrt{n}) \le 0.01$$

标准正态分布 $N(0,1)$ 的分布函数

$$\Phi(0.4\sqrt{n}) \ge 0.99$$

$n \ge 33.93$　所以n取34

单选题 1分

设置

作 H_0 、 H_1 的显著性检验，则下面说法正确的是

A 选择希望的结果为 H_0

B 选择临界域时只涉及 H_0

C 控制犯第一类错误的概率 α

D 使两类错误概率 α、 β 都很小

E 以上都不对

提交

实验目的

1、掌握参数估计和假设检验的基本理论模型与分类

2、根据问题的要求建立模型

3、对已经确定的模型，确定参数、使用MATLAB求解

作业

见网络学堂