概率统计第十五讲

假设检验

本讲题要

- 基本概念
- 单正态总体的假设检验
- 双正态总体均值差与方差比的假设检验

假设检验的基本概念和思想
（一）两类问题

1、参数假设检验

$X_1,\ldots,X_n \overset{\text{i.i.d.}}{\sim} F(x;\theta)$，$\theta\in\Theta$，总体分布已知，参数未知，由观测值$x_1$，…，$x_n$检验假设

$$H_0：\theta=\theta_0；\quad H_1：\theta\neq\theta_0。$$

2、非参数假设检验

$X_1,\ldots,X_n \overset{\text{i.i.d.}}{\sim} X$，总体分布未知，由观测值$x_1,\cdots,x_n$检验假设

$$H_0：F(x)=F_0(x;\theta)；\ H_1：F(x)\neq F_0(x;\theta)。$$

常称H_0为原假设，H_1为备择假设。

（二）检验法则与拒绝域

- 以样本$(X_1, \ldots, X_n)$出发制定一个法则，一旦观测值$(x_1, \ldots, x_n)$确定后，由这个法则就可作出判断是拒绝H_0还是接受H_0，这种法则称为H_0对H_1的一个检验法则，简称检验法。
- 样本观测值的全体组成样本空间S，把S分成两个互不相交的子集W和W^c，即
  $S=W\cup W^c$, $W\cap W^c=\varnothing$.
- 假设当$(x_1, \cdots, x_n)\in W$时，我们就拒绝H_0；当$(x_1, \cdots, x_n)\in W^c$时，我们就接受$H_0$。子集$W\subset S$就称为检验的拒绝域（或临界域）。一般将$W^c\subset S$称为接收域。

（三）检验的两类错误

称 H_0真而被拒绝的错误为**第一类错误**或**拒真错误**，

称 H_0假而被接受的错误为**第二类错误**或**受伪错误**。

记α=P(拒绝H_0| H_0真);

β=P(接受H_0| H_0假)。

对于给定的一对H_0和H_1, 总可找出许多拒绝域, 自然希望找到这种拒绝域W, 使得犯两类错误的概率都很小。

Neyman-Pearson提出了一个原则

“在控制犯第一类错误的概率不超过指定值α的条件下，尽量使犯第二类错误的概率β小”按这种法则做出的检验称为“显著性检验”，α称为显著性水平或检验水平。

如:对总体$X\sim N(\mu,1)$, 要检验

H_0: $\mu=0$; H_1: $\mu=1$。

拒绝域可取 $\bar{x}\geq k$，那么k=?

根据Neyman-Pearson原则：应选取k使“犯第一类错误的概率不超过指定值α的条件下,尽量使犯第二类错误的概率小”这里 $\mu=0$时，$\bar{x}\sim N(0,1/n)$.

$$\alpha(0)=P(\bar{x}\geq k\,|\,\mu=0)=1-\Phi\left(\sqrt{n}k\right)\leq\alpha,$$

$$\Rightarrow k\geq u_{1-\alpha}\big/\sqrt{n}.$$

而 $\mu=1$时，$\bar{x}\sim N(1,1/n)$.

$$\beta(1)=P(\bar{x}<k\,|\,\mu=1)=\Phi\left(\sqrt{n}(k-1)\right)$$

β关于***k***单增，故为使β小，***k***要尽可能小。

对比 $k\geq u_{1-\alpha}\big/\sqrt{n}$.

说明***k***最小只能取到 $u_{1-\alpha}\big/\sqrt{n}$，得水平为

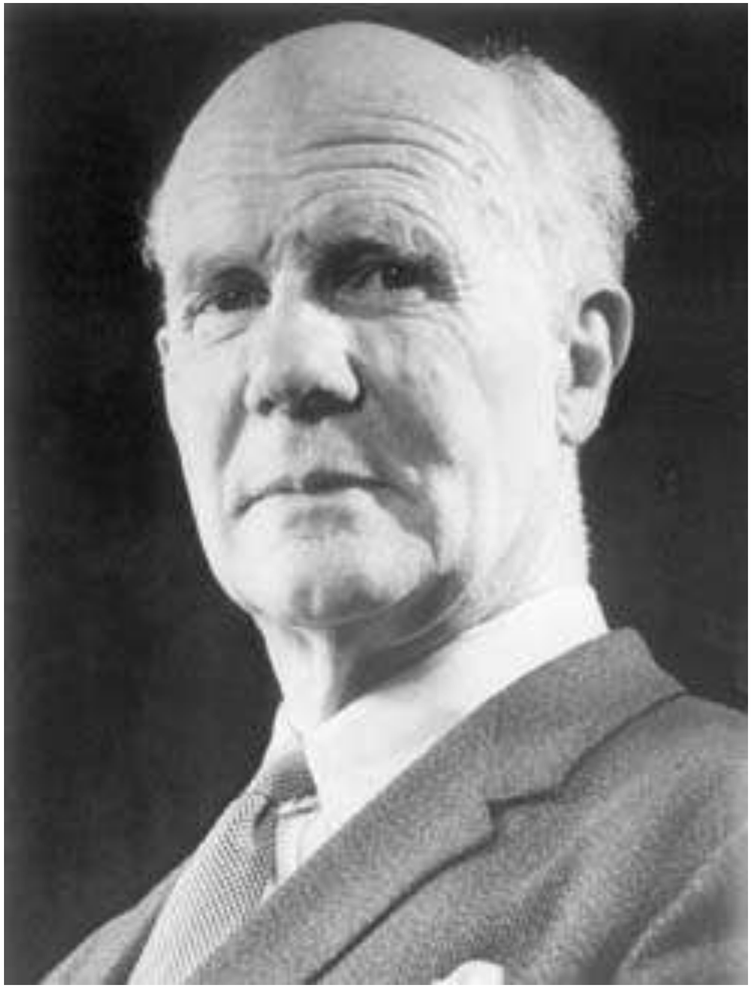
α的拒绝域为 $\bar{x}\geq u_{1-\alpha}\big/\sqrt{n}$.

- 可见，使α(0)≤α与又使β尽可能小的***k***值恰好满足α(0)=α。
- 一般地，符合Neyman-Pearson原则的拒绝域满足$\alpha(\theta_0)=\alpha$。

Jerzy Neyman

Born: 16 April 1894 in Bendery, Moldavia

Died: 5 Aug 1981 in Oakland, California, USA

Egon Sharpe Pearson

Born: 11 Aug 1895 in Hampstead (near London), England

Died: 12 June 1980 in Midhurst, Sussex, England

显著性检验的思想和步骤：

(1) 根据实际问题作出假设H_0与H_1；

(2) 构造统计量，在H_0真时其分布已知；

(3) 给定显著性水平α的值，参考H_1，令 $P(\text{拒绝}H_0 \mid H_0\text{真})=\alpha$，求出拒绝域$W$；

(4) 计算统计量的值，若统计量$\in W$，则拒绝H_0，否则接受H_0。

单正态总体的假设检验

一、单正态总体均值的假设检验

设$X_1,\cdots,X_n \overset{\text{i.i.d.}}{\sim} N(\mu,\sigma^2)$，给定检验水平α，由观测值 $x_1,\cdots,x_n$检验假设H_0：$\mu=\mu_0$；H_1：$\mu\neq\mu_0$。

1、σ已知的情形---u检验

对于假设H_0：$\mu=\mu_0$；H_1：$\mu\neq\mu_0$，构造

$$u=\frac{\sqrt{n}(\bar{x}-\mu)}{\sigma}\overset{H_0\text{真}}{=}\frac{\sqrt{n}(\bar{x}-\mu_0)}{\sigma}\sim N(0,1),$$

由$P(|u|\geq u_{1-\alpha/2})=\alpha$，可得拒绝域：$\{|u|\geq u_{1-\alpha/2}\}$.

查表，计算，比较大小，得出结论。

说明：

(1) H_0：$\mu=\mu_0$；H_1：$\mu\neq\mu_0$称为双边HT问题；

(2) H_0：$\mu\leq\mu_0$；H_1：$\mu>\mu_0$或H_0:$\mu\geq\mu_0$；H_1: $\mu<\mu_0$称为单边HT问题，这是完备的HT问题；

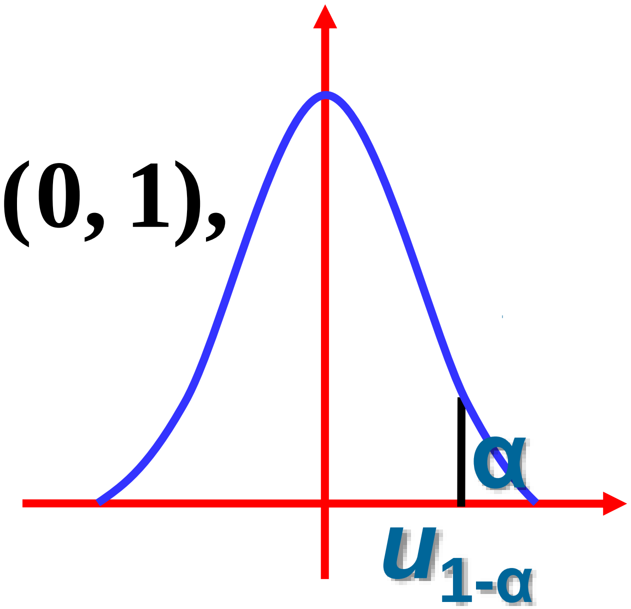
(3) H_0：$\mu=\mu_0$；H_1：$\mu>\mu_0$（或$\mu<\mu_0$），也称为单边问题，这是不完备的HT问题；

(4)可证：完备的HT问题与不完备的HT问题有相同的拒绝域，从而检验法一致。

不完备的右边HT问题的解

H_0: $\mu=\mu_0$; H_1: $\mu>\mu_0$。

$$H_0\text{下}，\ u=\frac{\sqrt{n}(\bar{x}-\mu_0)}{\sigma}\sim N(0,1),$$



由$P(u\geq u_{1-\alpha})=\alpha$，可得拒绝域：$W=\{u\geq u_{1-\alpha}\}$。

完备的右边HT问题

H_0：$\mu\le\mu_0$；H_1：$\mu>\mu_0$。

H_0下，$\dfrac{\sqrt{n}(\bar{x}-\mu)}{\sigma}\sim N(0,1)$，$\mu\le\mu_0$。令$u=\dfrac{\sqrt{n}(\bar{x}-\mu_0)}{\sigma}$，

若取拒绝域为$W=\{u\ge u_{1-\alpha}\}$（故这是合理的！）
则犯第一类错误的概率为

$$P(u\ge u_{1-\alpha}\mid\mu\le\mu_0)=P\left(\frac{\sqrt{n}(\bar{x}-\mu+\mu-\mu_0)}{\sigma}\ge u_{1-\alpha}\right)$$

$$=P\left(\frac{\sqrt{n}(\bar{x}-\mu)}{\sigma}\ge u_{1-\alpha}+\frac{\sqrt{n}(\mu_0-\mu)}{\sigma}\right)$$

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

于是，$\sup\limits_{\mu\le\mu_0}P(u\ge u_{1-\alpha}\mid\mu\le\mu_0)=\alpha.$

例1 设某厂生产一种灯管，其寿命$X\sim N(\mu,200^2)$，由以往经验知平均寿命$\mu=1500$小时，现采用新工艺后，在所生产的灯管中抽取25只，测得平均寿命1675小时，问采用新工艺后，灯管寿命是否有显著提高。（$\alpha=0.05$）

解：H_0：$\mu=1500$，H_1：$\mu>1500$。

$$H_0\text{下，}\ u=\frac{\sqrt{25}(\bar{x}-1500)}{200}\sim N(0,1)\text{。}$$

由$P(u\geq u_{1-\alpha})=\alpha$，可得拒绝域：$u\geq u_{0.95}=1.645$。

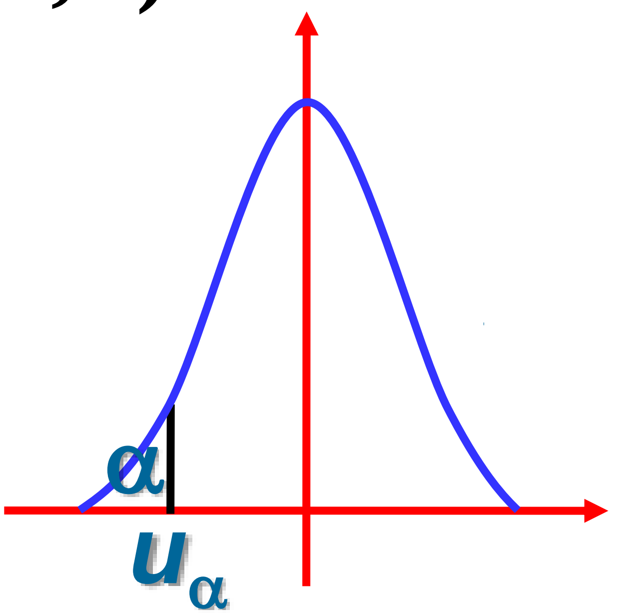
$$\text{这里}u=\frac{\sqrt{25}(1675-1500)}{200}=4.375>1.645,\ \text{拒绝}H_0\text{。}$$

左边HT问题

H_0：$\mu=\mu_0$；H_1：$\mu<\mu_0$或H_0：$\mu\geq\mu_0$；H_1：$\mu<\mu_0$。

$\mu=\mu_0$时，$\dfrac{\sqrt{n}(\bar{x}-\mu_0)}{\sigma}\sim N(0,1)$。

由$P(u\leq u_\alpha)=\alpha$，可得显著性水平为α的拒绝域为$u\leq u_\alpha$。



例2　已知某炼铁厂的铁水含碳量在正常情况下服从正态分布$N(4.55, 0.11^2)$. 某日测得5炉铁水含碳量如下：4.28, 4.40, 4.42, 4.35, 4.37. 如果标准差不变, 该日铁水的平均含碳量是否显著偏低？(取 $\alpha=0.05$)

解：H_0：$\mu=4.55$，H_1：$\mu<4.55$。

$$H_0\text{下}, \quad u=\frac{\sqrt{5}(\bar{x}-4.55)}{0.11} \sim N(0,1).$$

由$P(u \le u_\alpha)=\alpha$，得水平为α的拒绝域为 $u \le u_{0.05}=-1.645$。这里

$$u=\frac{\sqrt{5}(4.364-4.55)}{0.11}=-3.78<-1.645,$$

拒绝H_0。

注: 上题中, 用双边检验或右边检验都是错误的.

- 若用双边检验，H_0：$\mu=4.55$；H_1：$\mu\neq4.55$, 则拒绝域为$|u|\geq u_{1-\alpha/2}=1.96$。
- 由$|u|=3.78>1.96$, 故拒绝H_0, 说明可以认为该日铁水的平均含碳量显著异于4.55, 但无法说明是显著高于还是低于4.55. 不合题意。
- 若用右边检验，H_0：$\mu\leq4.55$；H_1：$\mu>4.55$, 则拒绝域为$u\geq u_{0.05}=-1.645$。
- 由$u=-3.78<-1.645$, 故接受H_0, 说明不能认为该日铁水的平均含碳量显著高于4.55, 但无法区分是等于还是低于4.55, 也不合题意。

2、σ未知的情形---t检验

双边检验：

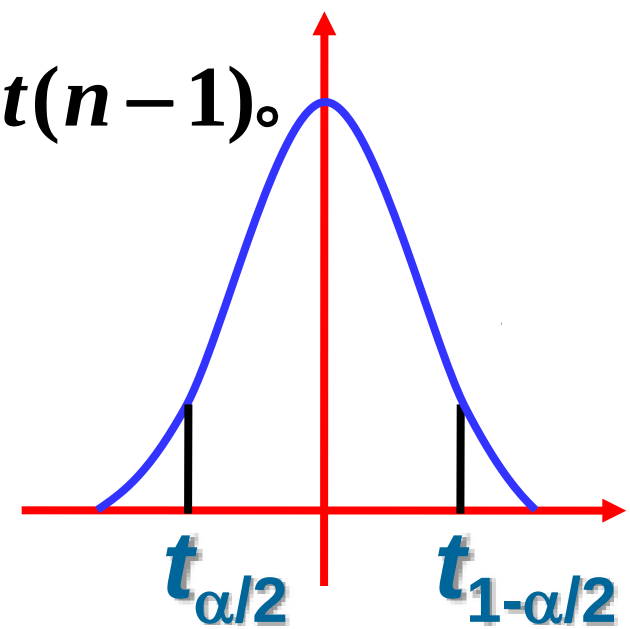
H_0： $\mu=\mu_0$； H_1： $\mu\neq\mu_0$。

$$H_0\text{真时：}\quad t=\frac{\sqrt{n}(\bar{x}-\mu_0)}{s}\sim t(n-1)。$$

由$P[|t|\geq t_{1-\alpha/2}(n-1)]=\alpha$,

得水平为α的拒绝域为

$|t|\geq t_{1-\alpha/2}(n-1)$。



例3 用热敏电阻测温仪间接测量地热勘探井底温度，重复测量7次，测得温度(℃)： 112.0 113.4 111.2 112.0 114.5 112.9 113.6 而用某种精确办法测得温度为112.6(可看作真值)，试问用热敏电阻测温仪间接测温有无系统偏差(设温度测量值X服从正态分布，取 $\alpha=0.05$)？

解：H_0：$\mu=112.6$；H_1：$\mu\neq112.6$。

$$H_0\text{真时：}\quad t=\frac{\sqrt{n}(\bar{x}-\mu_0)}{s}\sim t(n-1),$$

由$P[|t|\geq t_{0.975}(n-1)]=0.05$，得水平为$\alpha=0.05$的拒绝域为：$|t|\geq t_{0.975}(6)=2.4469$。这里

$$|t|=\left|\frac{\sqrt{7}(112.8-112.6)}{1.135}\right|=0.466<2.4469,$$ 接受H_0。

右边HT问题：

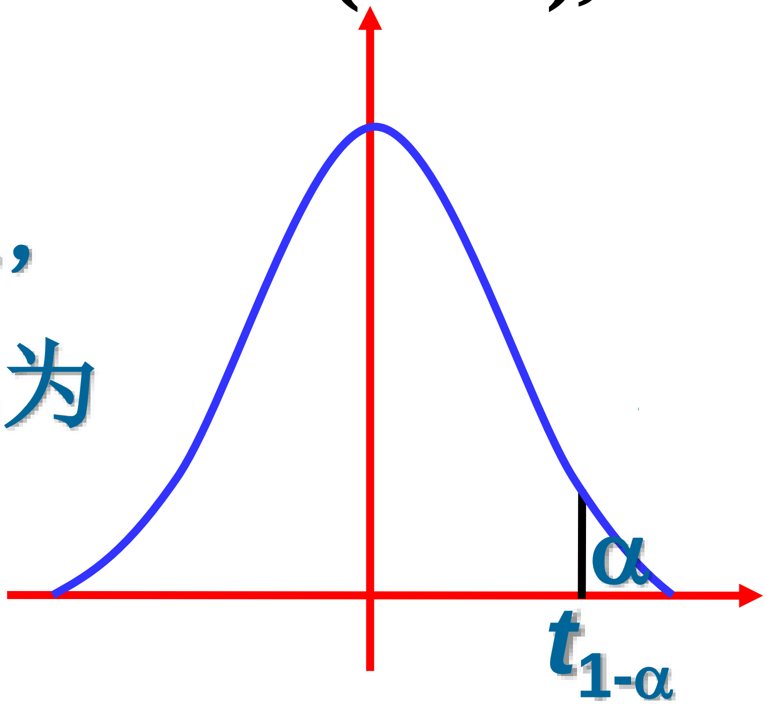
H_0:μ=μ_0； H_1:μ>μ_0或H_0:μ≤μ_0； H_1:μ>μ_0.

$$\mu=\mu_0\text{时：}\quad t=\frac{\sqrt{n}(\bar{x}-\mu_0)}{s}\sim t(n-1),$$

由$P[t\ge t_{1-\alpha}(n-1)]=\alpha$，

得水平为α的拒绝域为

$t\ge t_{1-\alpha}(n-1)$。



例4 某厂生产镍合金线，其抗拉强度的均值为10620 (kg/mm²) 今改进工艺后生产一批镍合金线，抽取10根, 测得抗拉强度 (kg/mm²) 为: 10512, 10623, 10668, 10554, 10776, 10707, 10557, 10581, 10666, 10670。认为抗拉强度服从正态分布, 取$\alpha=0.05$, 问新生产的镍合金线的抗拉强度是否比过去生产的合金线抗拉强度要高?

解：H_0：$\mu=10620$；H_1：$\mu>10620$。

$$H_0\text{真时：}\quad t=\frac{\sqrt{n}(\bar{x}-\mu_0)}{s}\sim t(n-1),$$

由$P[t\geq t_{0.95}(9)]=0.05$, 得拒绝域为$t\geq t_{0.95}(9)=1.8331$,

这里 $t=\dfrac{\sqrt{10}(10631.4-10620)}{81}=0.45<1.8331$, 接受$H_0$。

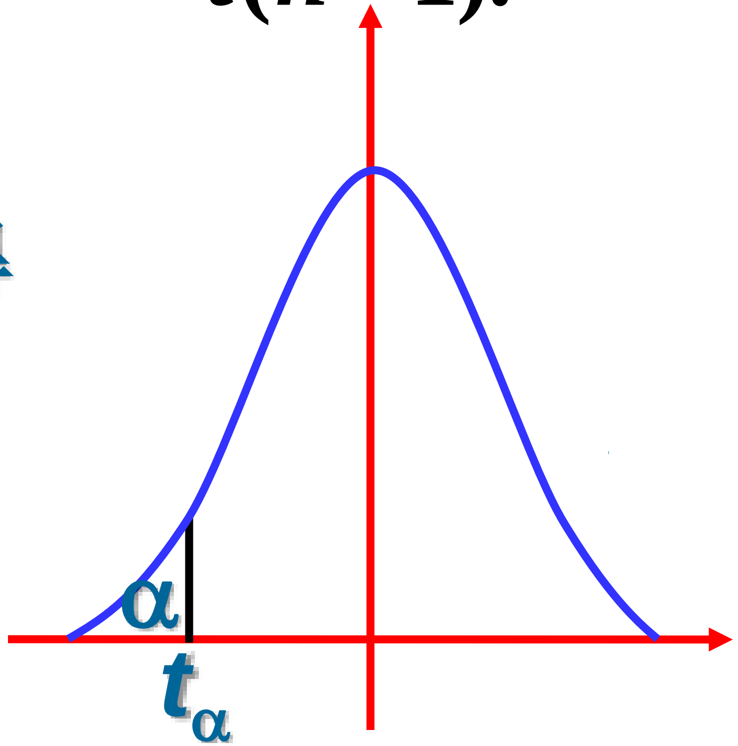
左边HT问题

H_0:μ=μ_0；H_1:μ<μ_0或H_0:μ≥μ_0；H_1:μ<μ_0。

$$\mu=\mu_0\text{时：}\quad t=\frac{\sqrt{n}(\bar{x}-\mu_0)}{s}\sim t(n-1).$$

由P[$t\le t_\alpha(n-1)$]=α，得水平为α的拒绝域为:

$t\le t_\alpha(n-1)$。



例5 设正品镍合金线的抗拉强度服从均值不低于10620(kg/mm^2)的正态分布，今从某厂生产的镍合金线中抽取10根, 测得平均抗拉强度10600 (kg/mm^2)，样本标准差为80. 问该厂的镍合金线的抗拉强度是否不合格？(α=0.1)

解：H_0：$\mu \geq 10620$；H_1：$\mu < 10620$。

$$\mu = 10620\text{时：}\quad t = \frac{\sqrt{10}(\bar{x} - 10620)}{s} \sim t(9),$$

由$P[t \leq t_{0.1}(9)] = 0.1$，得拒绝域为$t \leq t_{0.1}(9) = -1.383$.

这里 $t = \dfrac{\sqrt{10}(10600 - 10620)}{80} = -0.79 > -1.383$, 接受$H_0$。

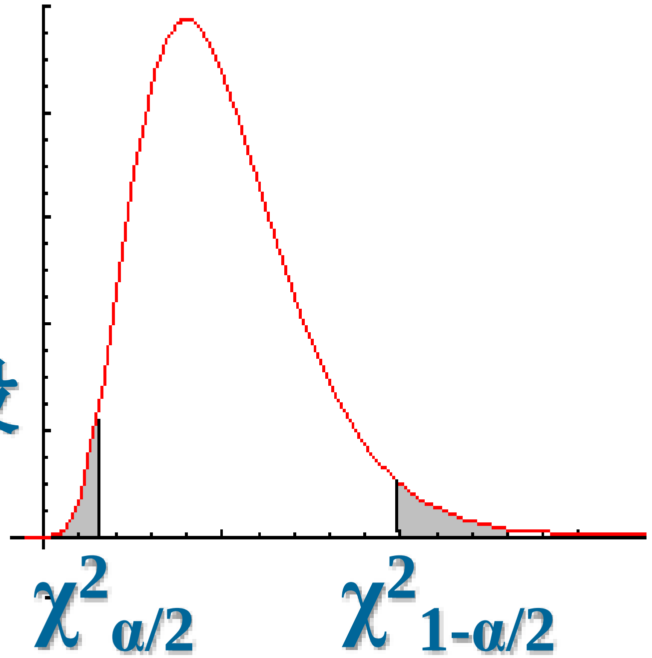
二、单总体方差的假设检验

设 $X_1,\cdots,X_n \overset{\text{i.i.d.}}{\sim} N(\mu,\sigma^2)$，给定检验水平α，由观测值 $x_1,x_2,\cdots,x_n$ 检验假设

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

假定μ未知，双边检验：在假设 H_0 下

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



由 $P[\chi^2 \le \chi^2_{\alpha/2}(n-1)$ 或 $\chi^2 \ge \chi^2_{1-\alpha/2}(n-1)]=\alpha$，得水平为α的拒绝域为

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

单边检验

对于右边问题：H_0：$\sigma^2=\sigma_0^2$， H_0：$\sigma^2>\sigma_0^2$，

可得拒绝域：$\chi^2 \geq \chi^2_{1-\alpha}(n-1)$;

而对左边问题：H_0：$\sigma^2=\sigma_0^2$， H_0：$\sigma^2<\sigma_0^2$，

可得拒绝域：$\chi^2 \leq \chi_\alpha^2(n-1)$。

例6 电工器材厂生产一批保险丝，取10根测得其熔化时间（min）为42，65，75，78，59，57，68，54，55，71. 问是否可以认为整批保险丝的熔化时间的方差小于等于80?(α=0.05, 熔化时间为正态变量。)

解：H_0: $\sigma^2 \le 80$； H_1: $\sigma^2 > 80$。这里 $\chi^2 = \dfrac{9s^2}{\sigma_0^2}$

$\sigma^2 = 80$时，$\chi^2 = \dfrac{9s^2}{80} \sim \chi^2(9)$，

由$P[\chi^2 \ge \chi^2_{1-\alpha}(9)] = \alpha$，

得水平为α=0.05的拒绝域为

$\chi^2 \ge \chi^2_{1-\alpha}(9) = \chi^2_{0.95}(9) = 16.919$。

$$= \frac{9 \times 121.8}{80} = 13.7$$

< 16.919，

接受H_0。

EX

设保险丝的融化时间服从正态分布，取9根测得其熔化时间（min）的样本均值为62，标准差为10.

(1) 是否可以认为整批保险丝的熔化时间服从$N(60, 9^2)$？（$\alpha=0.05$）

(2) 是否可以认为整批保险丝的熔化时间的方差显著大于70?（$\alpha=0.05$）

答：(1) $|t|=0.6<2.306$，接受60；$2.18<\chi^2=9.877<17.535$，接受9。

(2) $\chi^2=11.42<15.507$，认为方差不显著>70。

双正态总体均值差与方差比的假设检验

一、均值差的假设检验

设$X_1,\cdots,X_{n_1} \overset{\text{i.i.d.}}{\sim} N(\mu_1,\sigma_1^2)$; $Y_1,\cdots,Y_{n_2} \overset{\text{i.i.d.}}{\sim} N(\mu_2,\sigma_2^2)$,

两样本独立,给定检验水平 α，由观测值 $x_1,\cdots,x_{n_1}$;

$y_1,\cdots,y_{n_2}$检验假设 H_0： $\mu_1=\mu_2$; H_1： $\mu_1\neq\mu_2$。

1、σ_1，σ_2已知时的u检验

$$H_0\text{下，}\quad u=\frac{\bar{x}-\bar{y}}{\sqrt{\sigma_1^2/n_1+\sigma_2^2/n_2}}\sim N(0,1)。$$

2、$\sigma_1=\sigma_2=\sigma$但未知时的t检验

$$H_0\text{下，}\quad t=\frac{\bar{x}-\bar{y}}{s_w\sqrt{1/n_1+1/n_2}}\sim t(n_1+n_2-2)。$$

由$P(|u| \geq u_{1-\alpha/2}) = \alpha$，得拒绝域

$$|u| \geq u_{1-\alpha/2}。$$

由$P[|t| \geq t_{1-\alpha/2}(n_1+n_2-2)] = \alpha$，得拒绝域

$$|t| \geq t_{1-\alpha/2}(n_1+n_2-2)。$$

而对应的单边问题

H_0：$\mu_1=\mu_2$; H_1：$\mu_1>\mu_2$或H_0：$\mu_1\leq\mu_2$; H_1：$\mu_1>\mu_2$，拒绝域为 $u \geq u_{1-\alpha}$ 或 $t \geq t_{1-\alpha}(n_1+n_2-2)$；

H_0：$\mu_1=\mu_2$; H_1：$\mu_1<\mu_2$或H_0：$\mu_1\geq\mu_2$; H_1：$\mu_1<\mu_2$，拒绝域为 $u \leq u_{\alpha}$ 或 $t \leq t_{\alpha}(n_1+n_2-2)$。

例7. 比较甲, 乙两种安眠药的疗效。将20名患者分成两组, 每组10人. 其中10人服用甲药后延长睡眠的时数分别为1.9, 0.8, 1.1, 0.1, -0.1, 4.4, 5.5, 1.6, 4.6, 3.4; 另10人服用乙药后延长睡眠的时数分别为0.7, -1.6, -0.2, -1.2, -0.1, 3.4, 3.7, 0.8, 0.0, 2.0. 若服用两种安眠药后增加的睡眠时数服从方差相同的正态分布. 试问两种安眠药的疗效有无显著性差异?（$\alpha=0.10$）

解：H_0：$\mu_1=\mu_2$；H_1：$\mu_1\neq\mu_2$。

H_0下， $$t=\frac{\bar{x}-\bar{y}}{s_w\sqrt{1/10+1/10}}\sim t(18),$$

由$P[|t|\geq t_{0.95}(18)]=0.1$，

得拒绝域

$|t|\geq t_{0.95}(18)$

$=1.7341$。

这里：

$$\bar{x}=2.33,\quad s_1=2.002,\quad \bar{y}=0.75,\quad s_2=1.789,$$

$$s_w=\sqrt{\frac{9s_1^2+9s_2^2}{18}}\approx 1.898,$$

$$|t|=\frac{|\bar{x}-\bar{y}|}{s_w\sqrt{1/10+1/10}}\approx 1.86>1.7341,$$

拒绝H_0，认为两种安眠药的疗效有显著性差异。

EX1 上题中,试检验是否甲安眠药比乙安眠药疗效显著?

H_0：$\mu_1 \le \mu_2$；H_1：$\mu_1 > \mu_2$。

$$H_0\text{下},\quad t=\frac{\bar{x}-\bar{y}}{s_w\sqrt{1/10+1/10}}\sim t(18),$$

由$P[t \ge t_{0.9}(18)]=0.1$，得拒绝域

$t \ge t_{0.9}(18)=1.3304$。

这里：$t=1.86>1.3304$，故拒绝H_0，认为甲安眠药比乙安眠药疗效显著。

EX2 上题中,试检验是否乙安眠药比甲安眠药疗效显著?

二、方差比的假设检验

设$X_1,\cdots,X_{n_1} \overset{\text{i.i.d.}}{\sim} N(\mu_1,\sigma_1^2)$； $Y_1,\cdots,Y_{n_2} \overset{\text{i.i.d.}}{\sim} N(\mu_2,\sigma_2^2)$，

两样本独立，给定检验水平α，由观测值

$$x_1,\cdots,x_{n_1};\qquad y_1,\cdots,y_{n_2}$$

检验假设H_0: $\sigma_1^2=\sigma_2^2$； H_1: $\sigma_1^2\neq\sigma_2^2$。

假定μ_1，μ_2未知，则

H_0真时，$F:=s_1^2/s_2^2\sim F(n_1-1,n_2-1)$。

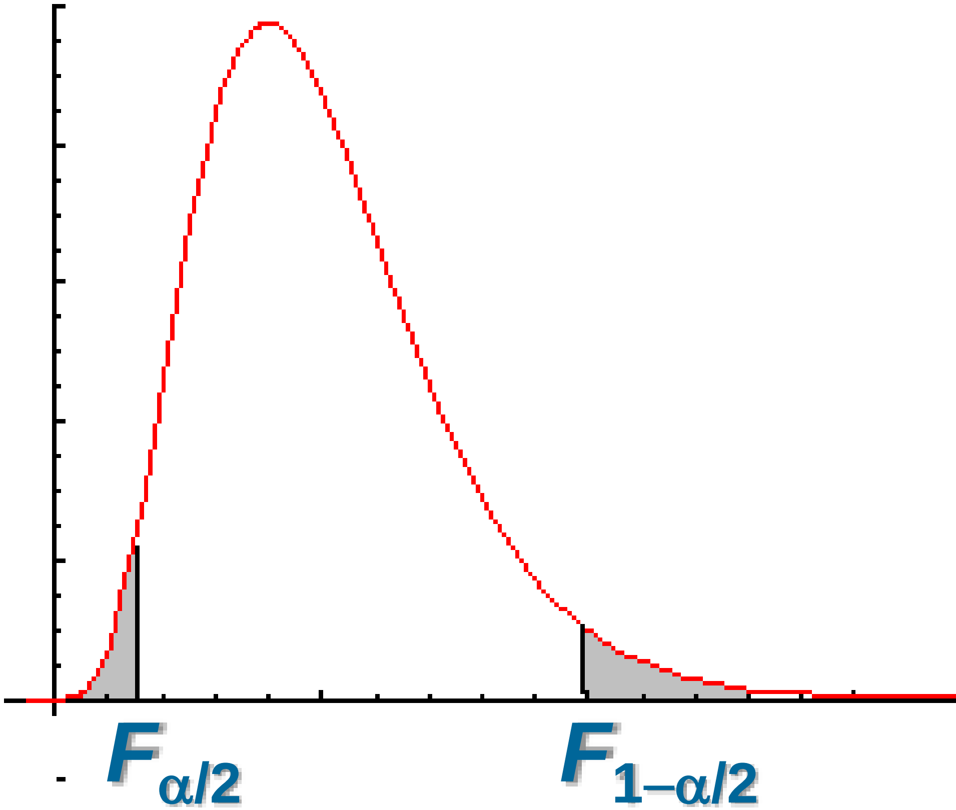
由$P[F \le F_{\alpha/2}(n_1-1, n_2-1)$或$F \ge F_{1-\alpha/2}(n_1-1, n_2-1)] = \alpha$,

得拒绝域

$F \le F_{\alpha/2}(n_1-1, n_2-1)$，

或

$F \ge F_{1-\alpha/2}(n_1-1, n_2-1)$。



而对应的单边问题

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

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

拒绝域为$F \ge F_{1-\alpha}(n_1-1, n_2-1)$;

$H_0: \sigma_1^2 = \sigma_2^2$; $H_1: \sigma_1^2 < \sigma_2^2$,

或$H_0: \sigma_1^2 \ge \sigma_2^2$; $H_1: \sigma_1^2 < \sigma_2^2$。

拒绝域为$F \le F_{\alpha}(n_1-1, n_2-1)$。

例8. 有甲乙两种机床, 加工同样产品, 从这两台机床加工的产品中 随 机 地 抽 取 若 干 产品, 测得产品直径为(单位:mm):

甲: 20.5, 19.8, 19.7, 20.4, 20.1, 20.9, 19.6, 19.9.

乙: 19.7, 20.8, 20.5, 19.8, 19.4, 20.6, 19.2.

假定甲, 乙 两台机床的产品直径都服从正态分布, 试比较甲, 乙两台机床加工的精度有无显著差异?($\alpha=0.05$)

解: H_0: $\sigma_1^2=\sigma_2^2$; H_1: $\sigma_1^2\neq\sigma_2^2$。

H_0真时, $F := s_1^2/s_2^2 \sim F(7,6)$。

拒绝域为$F \le F_{0.025}(7, 6)=1/5.12=0.1953$

或$F \ge F_{1-0.025}(7, 6)=F_{0.975}(7, 6)=5.7$。

这里: $s_1^2=0.204$, $s_2^2=0.397$, $F\approx0.51\in(0.1953,5.7)$,

故接受H_0。