第八章 假设检验

§1 假设检验

§2 正态总体均值和方差的假设检验

§1 假设检验

假设检验或“显著性检验（Test of statistical significance）”

- 是用来判断样本与样本，样本与总体的差异是由抽样误差引起还是本质差别造成的统计推断方法
- 先对总体的特征作出某种假设，然后通过抽样研究的统计推理，据此对假设做出拒绝或接受的推断

例：雾霾对动物有害？

将一些老鼠随机分成两组，其中一组置于洁净的空气中，另一组暴露于雾霾中

原假设/零假设：两组的肺病发病率相同

备择假设：两组的肺病发病率不相同

若暴露于雾霾中的一组的发病率远远高于另一组，则我们拒绝原假设，得到结论为：证据更支持备择假设

换句话说，证据表明雾霾会引起对动物的伤害

假设检验的基本思想

假设检验：事先对总体参数或分布形式做出某种假设，然后利用样本信息来判断原假设是否成立的过程

- 参数检验：在总体分布类型已知的前提下对总体参数及有关性质进行判断
- 非参数检验：总体分布的类型部分或全部未知，检验的目的是做出一般性的推断，例如分布的类型、两变量是否相互独立、分布是否相同等

问题1

已知某炼铁厂的铁水含碳量X在某种工艺条件下服从正态分布$N(4.55, 0.108^2)$。现在改变了工艺条件，又测了五炉铁水，其含碳量分别为：

4.28, 4.40, 4.42 4.35, 4.37

根据以往的经验，总体的方差$\sigma^2 = 0.108^2$一般不会改变。试问工艺条件改变后，铁水含碳量的均值有无改变？

需解决的问题：如何根据样本判断现在铁水的含碳量是服从$\mu \neq 4.55$的正态分布，还是仍然服从$\mu = 4.55$的正态分布。为此，选择其中之一作为假设，再利用样本检验假设的真伪

问题2

某工厂生产一批铁钉，从中随机抽取11根，测得长度（单位：mm）数据为：

$$10.41, 10.32, 10.62, 10.18, 10.77, 10.64,$$

$$10.82, 10.49, 10.38, 10.59, 10.54$$

试问铁钉的长度X是否服从正态分布？

需解决的问题：总体X是否服从正态分布。为此，选择是或否作为假设，然后利用样本对假设的真伪做出判断。

问题中涉及到的假设称为原假设，记为H_0；

与原假设对立的假设称为备择假设，记为H_1。

如问题1，若原假设为$H_0: \mu = \mu_0 = 4.55$, 则备择假设为$H_1: \mu \neq 4.55$。

如问题2，若原假设为$H_0: X$服从正态分布，则备择假设为$H_1: X$不服从正态分布。

假设检验的基本思想

一般采用反证法；结论的合理性依据统计学的小概率原理

在给定备择假设H_1下，利用样本对原假设H_0做出判断，若拒绝原假设H_0，则意味着接受备择假设H_1；否则，就不拒绝原假设H_0

小概率事件原理

- 小概率事件：统计学指在一次试验中，一个几乎不可能发生的概率称为小概率事件实际不可能原理
- 在一次试验中小概率事件一旦发生，原假设就是错误的

例1　已知某炼铁厂的铁水含碳量X在某种工艺条件下服从正态分布$N(4.55, 0.108^2)$。现在改变了工艺条件，又测了五炉铁水，其含碳量分别为：

4.28, 4.40, 4.42 4.35, 4.37

根据以往的经验，总体的方差$\sigma^2 = 0.108^2$一般不会改变。试问工艺条件改变后，铁水含碳量的均值有无改变？

解： 首先建立假设$H_0: \mu = \mu_0 = 4.55, H_1: \mu \neq 4.55$。

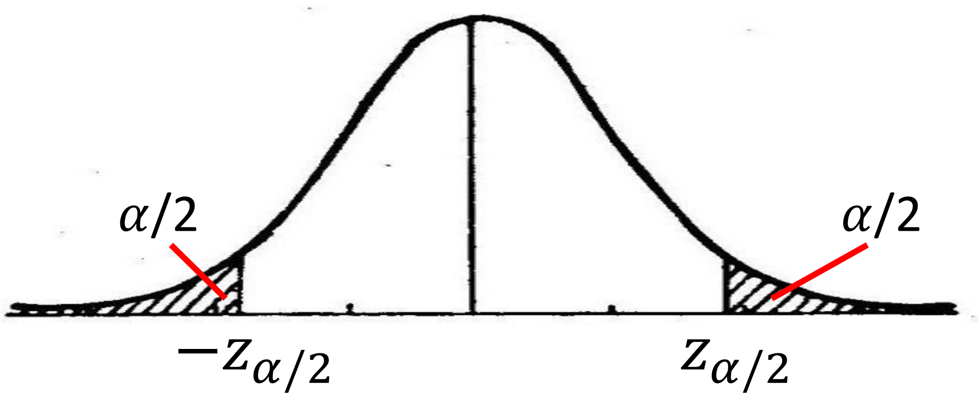
其次，从总体中随机抽样得到样本观察值$(x_1, x_2, \dots, x_n)$。

注意到$\bar{X} = \frac{1}{n}\sum_{i=1}^{n} x_i$是$\mu$的无偏估计量。因此，若$H_0$正确，则$\bar{X}$与$\mu_0$的偏差一般不应太大，即$|\bar{X} - \mu_0|$不应太大。由于$Z = \frac{\bar{X}-\mu_0}{\sigma/\sqrt{n}} \sim N(0,1)$，因此，考察$|\bar{X} - \mu_0|$的大小等价于考察$\frac{|\bar{X}-\mu_0|}{\sigma/\sqrt{n}}$的大小。

解（续）：

对于给定的小正数α，$P\left\{\frac{|\bar{X}-\mu_0|}{\sigma/\sqrt{n}} \geq z_{\alpha/2}\right\} = \alpha$。

事件$\left\{\frac{|\bar{X}-\mu_0|}{\sigma/\sqrt{n}} \geq z_{\alpha/2}\right\}$是小概率事件。



解（续）：

因此，当用样本值代入统计量$Z=\frac{\bar{X}-\mu_0}{\sigma/\sqrt{n}}$，具体计算得到其观察值$|z|=\frac{|\bar{X}-\mu_0|}{\sigma/\sqrt{n}}$时，

若$|z|\geq z_{\alpha/2}$，即说明在一次抽样中，小概率事件发生了。因此依据小概率原理，有理由拒绝H_0，接受H_1；

若$|z|<z_{\alpha/2}$，则没有理由拒绝H_0

统计量 Z 称为检验统计量。

当检验统计量取某个区域中的值时就拒绝，则称为的拒绝域，拒绝域的边界点称为临界点

在H_0和H_1之间，我们倾向于保护H_0 ，即H_0确实成立时，做出拒绝H_0的概率应是一个很小的正数，这类假设检验称为显著性假设检验，小正数 α 称为检验水平或显著性水平

参数的假设检验的一般步骤:

1. 根据实际问题建立原假设H_0和备择假设H_1；
2. 确定检验统计量Z，并明确其分布；
3. 给定显著性水平$\alpha > 0$，根据统计量的分布，由$P\{|Z| \geq z_{\alpha/2}\} \leq \alpha$确定拒绝域（临界值$z_{\alpha/2}$）；
4. 由样本值具体计算统计量Z的观察值z，并做出判断，若$|Z| \geq z_{\alpha/2}$，则拒绝H_0，接受H_1；若$|Z| < z_{\alpha/2}$，则不拒绝H_0

依此来解决例1提出的问题:

1. 假设$H_0: \mu = \mu_0 = 4.55$，$H_1: \mu \neq 4.55$ ；
2. 选择检验用统计量$z = \frac{\bar{X}-\mu_0}{\sigma/\sqrt{n}} \sim N(0,1)$ ；
3. 对于给定的小正数$\alpha = 0.05$，查标准正态分布表得到临界值$z_{\alpha/2} = z_{0.025} = 1.96$；
4. 具体计算：这里 $n = 5, \bar{x} = 4.364, \sigma^2 = 0.108^2$,故$Z$ 的观察值

$$z = \frac{\bar{X}-\mu_0}{\sigma/\sqrt{n}} = \frac{4.364-4.55}{0.108/\sqrt{5}} = -3.9$$

因为$|z| = 3.9 > 1.96$，所以拒绝H_0 ，接受H_1 ，即认为新工艺改变了铁水的平均含碳量。

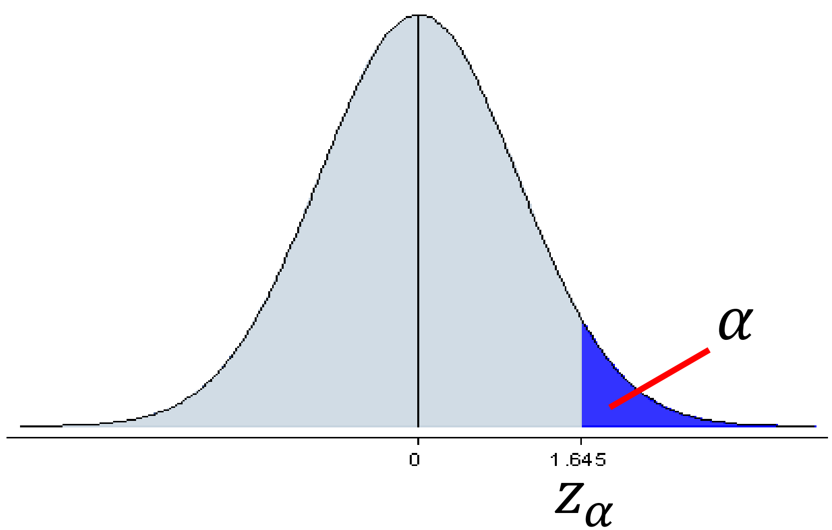
例 2 天然牛奶的冰点温度近似服从正态分布，均值$\mu_0 = -0.545℃$，标准差$\sigma = 0.008℃$。牛奶掺水可以使冰点温度升高至水的冰点温度（ 0℃ ）。测得生产商提交的 5 批牛奶的冰点温度，其均值为$\bar{x} = -0.535℃$，问是否可以认为生产商在牛奶中掺了水？取$\alpha = 0.05$。

解：设$H_0: \mu \le \mu_0$(未掺水)，$H_1: \mu > \mu_0$(已掺水)

解：设$H_0: \mu \le \mu_0$(未掺水)，$H_1: \mu > \mu_0$(已掺水)

这是单边检验问题，其拒绝域如下图所示，即为：

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



解（续）：

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

现有$z=\frac{-0.535-(-0.545)}{0.008/\sqrt{5}}=2.7951>1.645$。 z的值落在拒绝域中，所以在显著性水平$\alpha=0.05$下拒绝H_0，即认为牛奶商在牛奶中掺了水。

注意：

1. 原假设和备择假设地位不平等

假设检验是显著性检验。当显著性水平很小时检验推断“偏向”原假设。

在H_0和H_1之间，我们倾向于保护H_0，即H_0确实成立时，做出拒绝H_0的概率应是一个很小的正数，即显著性水平 α 。

对原假设的肯定缺乏说服力，但是否定原假设是有说服力的。

2. 显著性水平 α 是一个事先确定的小概率值。

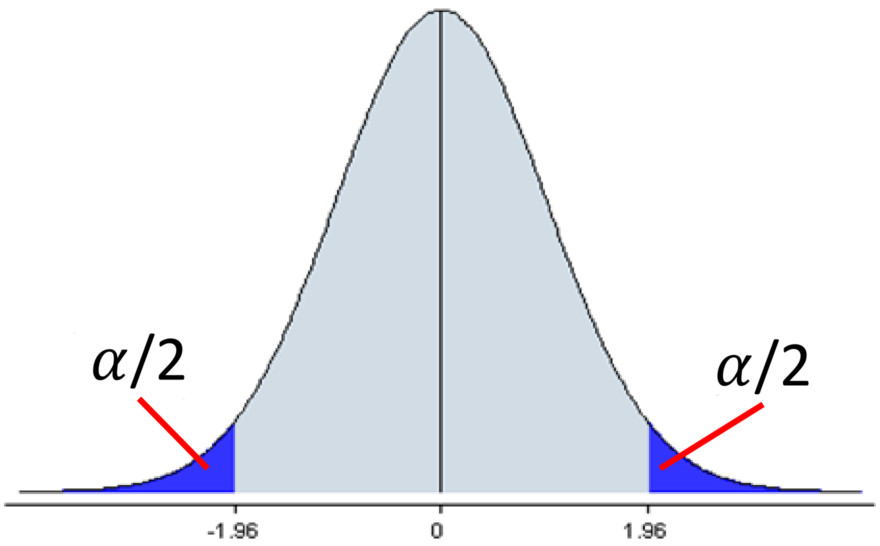
其取值事先确定，一般取$\alpha = 0.01$或$\alpha = 0.05$

3. 假设的形式分为单边检验和双边检验

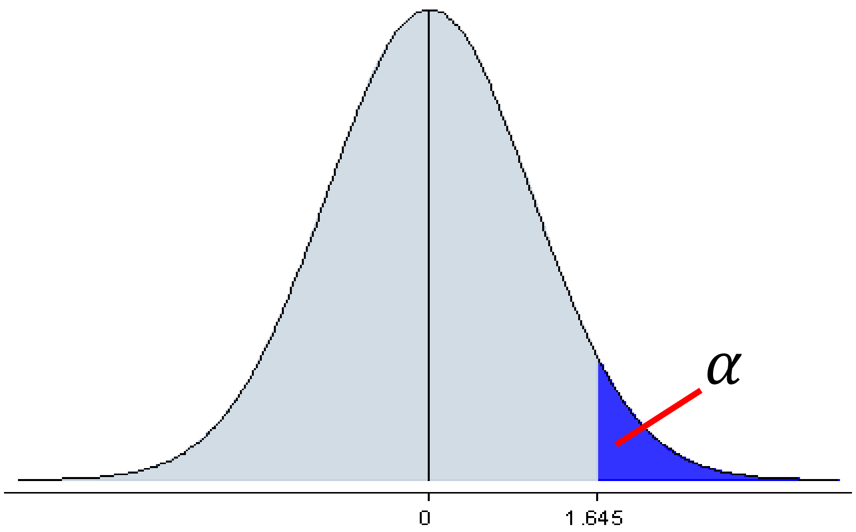
以方差已知，单个样本均值的显著性检验为例：

假设	研究的问题		
	双边检验	左边检验	右边检验
H_0	$\mu = \mu_0$	$\mu \geq \mu_0$	$\mu \leq \mu_0$
H_1	$\mu \neq \mu_0$	$\mu < \mu_0$	$\mu > \mu_0$

- 双边检验与单边检验



双边检验



右边检验

4. 如何选择统计量

明确要检验的是总体的什么参数

寻找与该参数近似的分布已知的统计量

5. 与区间估计的关系：原理相同，方法不同

原理：根据样本信息以抽样分布为理论依据在概率的基础上对总体参数进行推断，推断结果具有一定的可信程度和风险。

相关性：假设检验等价于检验原假设的值是否落在置信区间。若落在置信区间内，则不拒绝H_0，否则拒绝H_0。

方法：区间估计立足于大概率；假设检验立足于小概率

6. 假设检验中的两类错误

第 I 类错误：弃真错误，由于样本的随机性，犯这类错误是不可避免的。若将犯这一类错误的概率记为α，则有

$$P\{拒绝H_0|H_0为真\} = \alpha。$$

第 II 类错误：取伪错误，这类错误同样不可避免。若将犯这一类错误的概率记为β，则有

$$P\{接受H_0|H_0为假\} = \beta。$$

假设检验将犯第一类（弃真）错误的概率控制在给定的范围内，即显著性假设检验。

§2 正态总体均值和方差的检验

正态总体均值的检验

- 单个总体的均值检验
  - σ 已知时均值的检验： Z 检验
  - σ 未知时均值的检验： t 检验
- 两个总体的均值检验
  - $\sigma_1^2,\ \sigma_2^2$已知时均值的检验：Z 检验
  - $\sigma_1^2 = \sigma_2^2$未知时均值的检验：t 检验

正态总体方差的检验

- 单个总体方差的检验： χ^2检验
- 两个总体方差的检验： F 检验

再次回顾——三分布：

χ^2-分布

$X_1, X_2, \ldots, X_n$是来自总体$N(0,1)$的样本，则

$$\chi^2 = X_1^2 + X_2^2 + \cdots + X_n^2 \sim \chi^2(n)$$

性质：

$E(\chi^2) = n$，$D(\chi^2) = 2n$，$\chi_1^2 + \chi_2^2 \sim \chi^2(n_1 + n_2)$

$Student's\ t$-分布

$X \sim N(0,1), Y \sim \chi^2(n)$，且$X$、$Y$相互独立，则

$$t = \frac{X}{\sqrt{Y/n}} \sim t(n)$$

再次回顾——三分布：

F-分布

$U \sim \chi^2(n_1), V \sim \chi^2(n_2)$，且$U$、$V$相互独立，则

$$F = \frac{U/n_1}{V/n_2} \sim F(n_1, n_2)$$

F-分布的上α分位点的性质：

$$F_{1-\alpha}(n_1, n_2) = \frac{1}{F_\alpha(n_2, n_1)}$$

再次回顾——正态总体的抽样四定理：

定理1. 设$X_1, X_2, \dots, X_n$ 是来自总体$X \sim N(\mu, \sigma^2)$的样本，则

$$\bar{X} \sim N(\mu, \sigma^2/n)$$

定理2. 设$X_1, X_2, \dots, X_n$ 是来自总体$X \sim N(\mu, \sigma^2)$的样本，$\bar{X}$和S^2分别为样本均值与样本方差，则有

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

定理3. 设$X_1, X_2, \dots, X_n$ 是来自总体$X \sim N(\mu, \sigma^2)$的样本，$\bar{X}$和S^2分别为样本的均值与样本方差，则有

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

再次回顾——正态总体的抽样四定理：

定理4. 设$X_1, X_2, \dots, X_n$ 与$Y_1, Y_2, \dots, Y_n$分别是来自总体$X \sim N(\mu_1, \sigma_1^2)$和$Y \sim N(\mu_2, \sigma_2^2)$的样本，且两样本分别独立，两样本均值与样本方差分别为$\bar{X}, \bar{Y}, S_1^2, S_2^2$ ，则有

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

当$\sigma_1^2 = \sigma_2^2 = \sigma^2$时，

$$\frac{(\bar{X}-\bar{Y})-(\mu_1-\mu_2)}{S_w\sqrt{1/n_1+1/n_2}} \sim t(n_1+n_2-2),$$

其中

$$S_w^2 = \frac{(n_1-1)S_1^2+(n_2-1)S_2^2}{n_1+n_2-2}, \qquad S_w = \sqrt{S_w^2}$$

- 单个正态总体均值的检验

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

(1) σ^2已知

在H_0成立的条件下，选用检验统计量

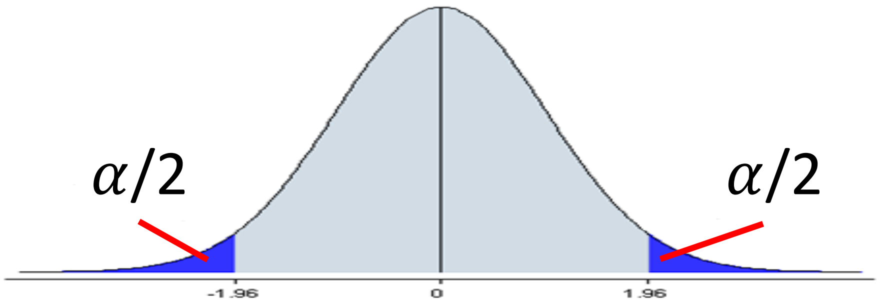
$$Z = \frac{\bar{X} - \mu_0}{\sigma/\sqrt{n}} \sim N(0,1)$$

对于给定的检验水平α，查正态分布表得到临界值$z_{\alpha/2}$，再由样本值具体计算统计量Z的观察值 z 并与$z_{\alpha/2}$比较。

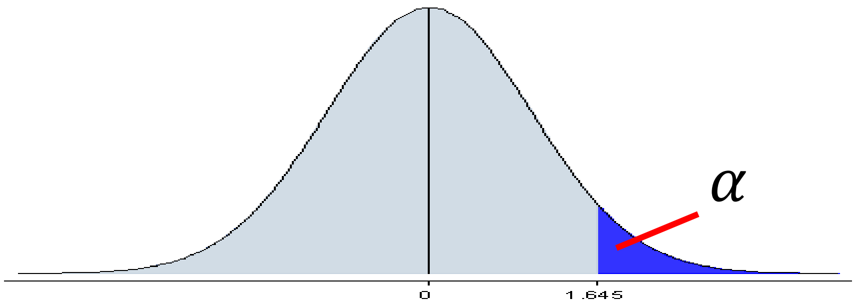
若$|z| \geq z_{\alpha/2}$，则拒绝H_0 ，接受H_1 ；若$|z| < z_{\alpha/2}$，则接受H_0 。这种检验法称为Z检验法。

不同原假设和备择假设形式下的拒绝域示意图

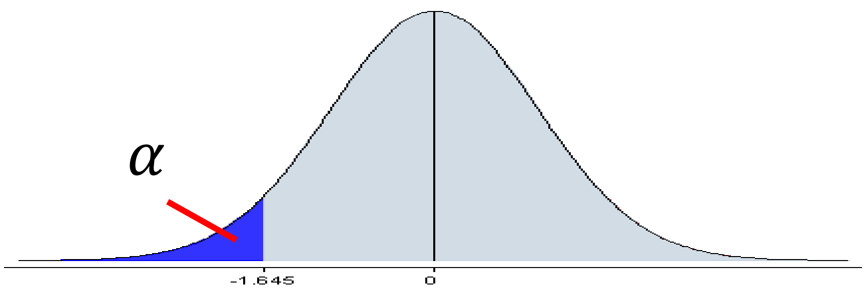
$H_0: \mu = \mu_0, H_1: \mu \neq \mu_0$



$H_0: \mu \leq \mu_0, H_1: \mu > \mu_0$



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



(2) σ^2未知

样本方差 $S^2 = \frac{1}{n-1}\sum_{i=1}^{n}(X_i - \bar{X})^2$ 是 σ^2 的无偏估计，因此用S^2代替σ^2。

在H_0成立的条件下，统计量 $T = \frac{\bar{X}-\mu_0}{S/\sqrt{n}} \sim t(n-1)$。

给定显著性水平$\alpha > 0$，查 t 分布表可得临界值$t_{\alpha/2}$,使$P\{|t| \geq t_{\alpha/2}\} = \alpha$ 成立。再由样本值具体计算统计量T的观察值t,并与$t_{\alpha/2}$比较，若$|t| \geq t_{\alpha/2}$，则拒绝H_0 ，接受H_1 ；若$|t| < t_{\alpha/2}$，则接受H_0 。这种检验法称为 t检验法。

例 3 某元器件的寿命X（以 h 计）服从正态分布$N(\mu,\sigma^2)$，μ,σ^2均未知。现测得16只元器件寿命如下：

$$159, 280, 101, 212, 224, 379, 179, 264$$

$$222, 362, 168, 250, 149, 260, 485, 170$$

问：是否有理由认为元件的平均寿命大于225 h。

解： $H_0: \mu \leq \mu_0 = 225, H_1: \mu > 225$

检验问题的拒绝域为：

$$t = \frac{\bar{x} - \mu_0}{s/\sqrt{n}} \geq t_\alpha(n-1)$$

解（续）：检验问题的拒绝域为：

$$t=\frac{\bar{x}-\mu_0}{s/\sqrt{n}}\geq t_\alpha(n-1)$$

取$\alpha=0.05$，$t_{0.05}(15)=1.7531$。

由样本算得：$\bar{x}=241.5$，$s=98.7259$，即有

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

t没有落在拒绝域中，故接受H_0，即认为元器件的平均寿命不大于225 h。

- 两个正态总体均值差的检验

(1) σ_1^2、σ_2^2 已知

原假设H_0: $\mu_1 - \mu_2 = \delta$ ，备择假设H_1: $\mu_1 - \mu_2 \neq \delta$

在H_0成立的条件下，检验统计量

$$Z = \frac{\bar{X} - \bar{Y} - \delta}{\sqrt{\frac{\sigma_1^2}{n_1} + \frac{\sigma_2^2}{n_2}}} \sim N(0,1)$$

(2) $\sigma_1^2=\sigma_2^2$ 未知

原假设$H_0\colon \mu_1-\mu_2=\delta$，备择假设$H_1\colon \mu_1-\mu_2\neq\delta$

在H_0成立的条件下，检验统计量

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

其中，$S_w^2=\frac{(n_1-1)S_1^2+(n_2-1)S_2^2}{n_1+n_2-2}$。

对于给定的检验水平α，查 t 分布表得到临界值 $t_{\frac{\alpha}{2}}(n_1+n_2-2)$，再由样本值具体计算统计量 T 的观察值t。

若$|t| \geq t_{\frac{\alpha}{2}}(n_1+n_2-2)$，则拒绝$H_0$ ，接受H_1 ；若 $|t| < t_{\frac{\alpha}{2}}(n_1+n_2-2)$，则接受$H_0$ 。这种检验法称为 t 检验法。

例 4 用方法A和B测定冰自-0.72°C转变为0°C的水的融化热（以cal/g计）如下：

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: $80.02, 79.94, 79.98, 79.97, 79.97, 80.03, 79.95, 79.97$

设样本相互独立，且服从正态分布$N(\mu_1, \sigma^2)$和$N(\mu_2, \sigma^2)$，μ_1, μ_2，σ^2均未知。试比较两组测量结果的均值（取$\alpha = 0.05$），即检验假设：

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

解：方法A和B的箱线图如下：

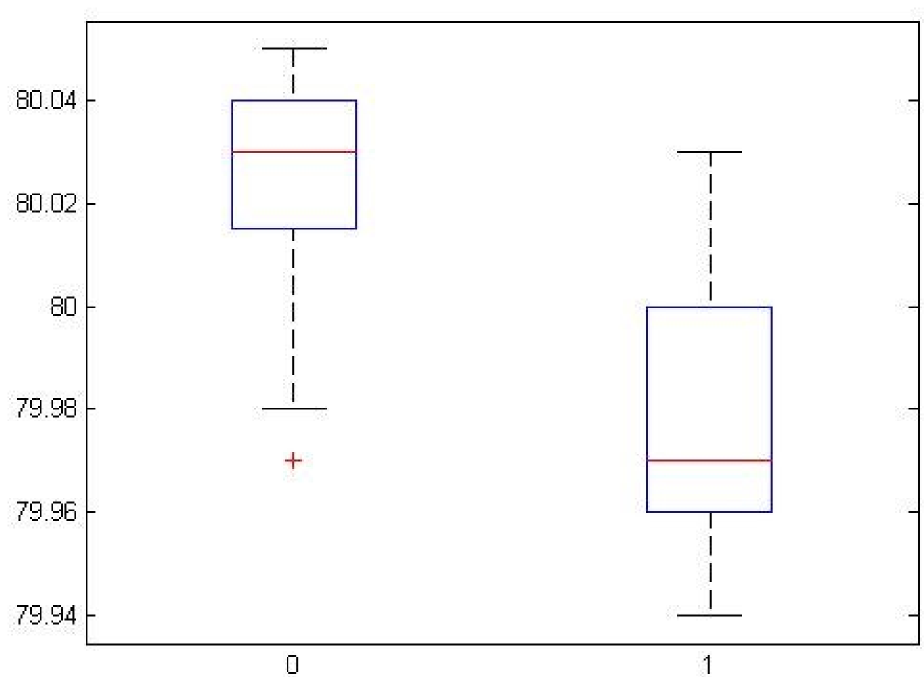
由样本计算得：

$n_1 = 13, \bar{x}_A = 80.02, s_A^2 = 0.024^2$

$n_2 = 8, \bar{x}_B = 79.98, s_B^2 = 0.031^2$

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

$$= 0.0007178$$



$$S_w^2 = \frac{(n_1 - 1)S_1^2 + (n_2 - 1\ )S_2^2}{n_1 + n_2 - 2}$$

解（续）：

$$t=\frac{\bar{x}_A-\bar{x}_B}{s_w\sqrt{1/13+1/8}}=3.333>t_{0.05}(13+8-2)$$

$$=1.7291$$

故拒绝H_0，认为方法A比方法B测得的融化热要大。

- 成对数据的均值差的检验

n对相互独立的样本: $(X_1, Y_1), (X_2, Y_2), \dots, (X_n, Y_n)$, $D_i = X_i - Y_i \sim N(\mu_D, \sigma_D^2)$。

原假设$H_0: \mu_D = 0$ ，备择假设$H_1: \mu_D \neq 0$

样本均值和方差记作：$\bar{D}, S_D^2$

在H_0成立的条件下，检验统计量

$$\frac{\bar{D}}{S_D/\sqrt{n}} \sim t(n-1)$$

例 5 用两台光谱仪I_x和I_y测量9件材料试块中的金属含量，得到以下9对观察值：

$x(\%)$	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

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：设$D = X - Y$，则$D_1, D_2, \dots, D_n$独立并服从同一分布。假设$D_i \sim N(\mu_D, \sigma_D^2), i = 1, \dots, n$。

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

解（续）：检验统计量$\frac{\bar{D}}{S_D/\sqrt{n}} \sim t(n-1)$

由$n=9, t_{0.005}(8)=3.3554$可算得拒绝域为

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

由样本观察值得$\bar{d}=0.06, s_D=0.122, |t|=1.467 < 3.3554$。

$|t|$值没有落在拒绝域中，故接受H_0，即认为两台仪器的测量结果无明显差异。

- 单个正态总体方差的检验

(1) μ 已知

原假设H_0: $\sigma^2 = \sigma_0^2$ ，备择假设H_1: $\sigma^2 \neq \sigma_0^2$

检验统计量

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

(2) μ 未知

原假设$H_0: \sigma^2 = \sigma_0^2$ ，备择假设$H_1: \sigma^2 \neq \sigma_0^2$

$S^2 = \frac{1}{n-1}\sum_{i=1}^{n}(X_i - \bar{X})^2$是$\sigma^2$的无偏估计量。

在H_0成立的条件下，统计量

$$\chi^2 = \frac{\sum_{i=1}^{n}(X_i - \bar{X})^2}{\sigma_0^2} = \frac{(n-1)S^2}{\sigma_0^2} \sim \chi^2(n-1)$$

对于给定的检验水平α，查χ^2分布表得到临界值$\chi^2_{\frac{\alpha}{2}}(n-1)$及$\chi^2_{1-\frac{\alpha}{2}}(n-1)$，使得

$$P\left\{\chi^2 \geq \chi^2_{\frac{\alpha}{2}}(n-1)\right\} = P\left\{\chi^2 \leq \ \chi^2_{1-\frac{\alpha}{2}}(n-1)\right\} = \frac{\alpha}{2}$$

再由样本值具体计算统计量χ^2的观察值$\frac{(n-1)S^2}{\sigma_0^2}$

若$\frac{(n-1)S^2}{\sigma_0^2} \leq \chi^2_{1-\frac{\alpha}{2}}(n-1)$或$\frac{(n-1)S^2}{\sigma_0^2} \geq \chi^2_{\alpha/2}(n-1)$，则拒绝$H_0$，接受$H_1$；否则接受$H_0$。这种检验法称为$\chi^2$检验法。

注意：对于给定的检验水平α，单个正态总体的方差检验也可能是单边检验问题。

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

右边检验的拒绝域为：$\frac{(n-1)S^2}{\sigma_0^2} \ge \chi_\alpha^2(n-1)$

$H_0: \sigma^2 \ge \sigma_0^2$ ，$H_1: \sigma^2 < \sigma_0^2$

左边检验的拒绝域为：$\frac{(n-1)S^2}{\sigma_0^2} \le \chi_{1-\alpha}^2(n-1)$

- 两个正态总体方差的检验

(1) μ_1, μ_2 已知

原假设$H_0: \sigma_1^2 = \sigma_2^2$ ，备择假设$H_1: \sigma_1^2 \neq \sigma_2^2$

在H_0成立的条件下，检验统计量

$$F = \frac{\frac{1}{n_1}\sum_{i=1}^{n_1}(X_i - \mu_1)^2}{\frac{1}{n_2}\sum_{i=1}^{n_2}(X_i - \mu_2)^2} \sim F(n_1, n_2)$$

(2) μ_1, μ_2 未知

原假设$H_0: \sigma_1^2 = \sigma_2^2$ ，备择假设$H_1: \sigma_1^2 \neq \sigma_2^2$

在H_0成立的条件下，检验统计量

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

对于给定的检验水平α，由

$$P\left\{F \geq F_{\frac{\alpha}{2}}(n_1-1, n_2-1)\right\} = P\left\{F \leq F_{1-\frac{\alpha}{2}}(n_1-1, n_2-1)\right\} = \frac{\alpha}{2}$$

可以得到临界值$F_{\frac{\alpha}{2}}(n_1-1, n_2-1)$和$F_{1-\frac{\alpha}{2}}(n_1-1, n_2-1)$，再由样本值具体计算统计量 F 的观察值f。

若$f \geq F_{\frac{\alpha}{2}}(n_1-1, n_2-1)$或$f \leq F_{1-\frac{\alpha}{2}}(n_1-1, n_2-1)$，则拒绝$H_0$，接受$H_1$；否则接受$H_0$。这种检验法称为$F$检验法。

单边检验：

原假设H_0: $\sigma_1^2 \leq \sigma_2^2$ ，备择假设H_1: $\sigma_1^2 > \sigma_2^2$

在H_0成立的条件下，$E(S_1^2) = \sigma_1^2 \leq \sigma_2^2 = E(S_2^2)$。

检验统计量$F = \frac{S_1^2/S_2^2}{\sigma_1^2/\sigma_2^2} \sim F(n_1 - 1, n_2 - 1)$。 由于

$P\{F \geq F_\alpha(n_1 - 1, n_2 - 1)\} = \alpha$ ，$\frac{S_1^2/S_2^2}{\sigma_1^2/\sigma_2^2} \geq \frac{S_1^2}{S_2^2}$ ，

所以拒绝域为：$\frac{S_1^2}{S_2^2} \geq F_\alpha(n_1 - 1, n_2 - 1)$

例 6 用方法A和B测定冰自$-0.72°C$转变为$0°C$的水的融化热（以cal/g计）如下：

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: $80.02, 79.94, 79.98, 79.97, 79.97, 80.03,$

$79.95, 79.97$

设样本相互独立，且服从正态分布$N(\mu_A, \sigma_A^2)$和$N(\mu_B, \sigma_B^2)$，试检验假设$H_0: \sigma_A^2 = \sigma_B^2$，$H_1: \sigma_A^2 \neq \sigma_B^2$(取$\alpha = 0.01$)。

$$F_{1-\alpha}(n_1,n_2)=\frac{1}{F_\alpha(n_2,n_1)}$$

解：此处，

拒绝域为$\frac{s_A^2}{s_B^2}\geq F_{0.005}(12,7)=8.18$或

$$\frac{s_A^2}{s_B^2}\leq F_{0.995}(12,7)=\frac{1}{F_{0.005}(7,12)}=\frac{1}{5.52}=0.18$$

由 $n_1=13,\bar{x}_A=80.02,s_A^2=0.024^2$

$n_2=8,\bar{x}_B=79.98,s_B^2=0.031^2,\frac{s_A^2}{s_B^2}=0.64$

$0.18<0.64<8.18$，故接受H_0，即认为两总体方差相等。

小结：正态总体均值、方差检验法

（1）检验统计量的构造及其分布与原假设H_0有关，均值用 Z 检验（ σ^2已知）或 t 检验（ σ^2未知）；方差检验用 χ^2检验（单总体）或 F 检验（两个总体）。

（2）拒绝域的选取与备择假设H_1有关，应该在对备择假设有利的方向上选取拒绝域（单边或双边）。

正态总体均值检验（单个总体）

	原假设H_0	备择假设H_1	检验统计量	拒绝域
1	$\mu \le \mu_0$ $\mu \ge \mu_0$ $\mu = \mu_0$ (σ^2已知)	$\mu > \mu_0$ $\mu < \mu_0$ $\mu \ne \mu_0$	$Z = \dfrac{\bar{X} - \mu_0}{\sigma/\sqrt{n}}$	$z \ge z_\alpha$ $z \le -z_\alpha$ $\lvert z \rvert \ge z_{\alpha/2}$
2	$\mu \le \mu_0$ $\mu \ge \mu_0$ $\mu = \mu_0$ (σ^2未知)	$\mu > \mu_0$ $\mu < \mu_0$ $\mu \ne \mu_0$	$t = \dfrac{\bar{X} - \mu_0}{S/\sqrt{n}}$	$t \ge t_\alpha(n-1)$ $t \le -t_\alpha(n-1)$ $\lvert t \rvert \ge t_{\alpha/2}(n-1)$

正态总体均值检验（两个总体、成对数据）

	原假设H_0	备择假设H_1	检验统计量	拒绝域
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)	$\mu_1-\mu_2>\delta$ $\mu_1-\mu_2<\delta$ $\mu_1-\mu_2\ne\delta$	$Z=\dfrac{\bar{X}-\bar{Y}-\delta}{\sqrt{\dfrac{\sigma_1^2}{n_1}+\dfrac{\sigma_2^2}{n_2}}}$	$z\ge z_\alpha$ $z\le -z_\alpha$ $\lvert z\rvert\ge z_{\alpha/2}$
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)	$\mu_1-\mu_2>\delta$ $\mu_1-\mu_2<\delta$ $\mu_1-\mu_2\ne\delta$	$T=\dfrac{\bar{X}-\bar{Y}-\delta}{S_w\sqrt{\dfrac{1}{n_1}+\dfrac{1}{n_2}}}$ $S_w^2=\dfrac{(n_1-1)S_1^2+(n_2-1)S_2^2}{n_1+n_2-2}$	$t\ge t_\alpha(n_1+n_2-2)$ $t\le -t_\alpha(n_1+n_2-2)$ $\lvert t\rvert\ge t_{\frac{\alpha}{2}}(n_1+n_2-2)$
5	$\mu_D\le 0$ $\mu_D\ge 0$ $\mu_D=0$ (成对数据)	$\mu_D>0$ $\mu_D<0$ $\mu_D\ne 0$	$t=\dfrac{\bar{D}-0}{S_D/\sqrt{n}}$	$t\ge t_\alpha(n-1)$ $t\le -t_\alpha(n-1)$ $\lvert t\rvert\ge t_{\alpha/2}(n-1)$

正态总体方差检验（单个、两个总体）

	原假设H_0	备择假设H_1	检验统计量	拒绝域
6	$\sigma^2 \le \sigma_0^2$ $\sigma^2 \ge \sigma_0^2$ $\sigma^2 = \sigma_0^2$ (μ未知)	$\sigma^2 > \sigma_0^2$ $\sigma^2 < \sigma_0^2$ $\sigma^2 \ne \sigma_0^2$	$\chi^2 = \dfrac{(n-1)S^2}{\sigma_0^2}$	$\chi^2 \ge \chi_\alpha^2(n-1)$ $\chi^2 \le \chi_{1-\alpha}^2(n-1)$ $\chi^2 \ge \chi_{\alpha/2}^2(n-1)$或 $\chi^2 \le \chi_{1-\alpha/2}^2(n-1)$
7	$\sigma_1^2 \le \sigma_2^2$ $\sigma_1^2 \ge \sigma_2^2$ $\sigma_1^2 = \sigma_2^2$ (μ_1,μ_2未知)	$\sigma_1^2 > \sigma_2^2$ $\sigma_1^2 < \sigma_2^2$ $\sigma_1^2 \ne \sigma_2^2$	$F = \dfrac{S_1^2}{S_2^2}$	$F \ge F_\alpha(n_1-1, n_2-1)$ $F \le F_{1-\alpha}(n_1-1, n_2-1)$ $F \ge F_{\alpha/2}(n_1-1, n_2-1)$或 $F \le F_{1-\alpha/2}(n_1-1, n_2-1)$

注意：

(1) 对于H_0只能说拒绝与不拒绝，而对H_1只能说接受。

(2) $P \le \alpha$，则拒绝H_0，接受H_1，差异有统计学意义，即有足够的证据可认为……不同或不等。

(3) $P > \alpha$，则不拒绝H_0，差距无统计学意义，尚不能认为……不同或不等。

(4) 做统计检验结论时只能说有无统计学意义，而不能说明专业上的差异大小。P值越小只能说明拒绝H_0的统计学证据越充分，推论时犯错误的机会越小，与专业上$|\mu-\mu_0|$差异的大小无直接关系。

(5) 事先确定显著性水平α。选择$\alpha=0.05$只是一种习惯，而不是绝对的标准。

作 业

《概率论与数理统计》：

P. 219, #6，#8，#10, #17