第八章　假设检验

关键词：

假设检验

正态总体参数的假设检验

拟合优度检验

§8.1 假设检验的基本思想

假设检验是数理统计的一类基本而重要的问题，特别在质量控制中有普遍应用。假设检验包括：

（1）已知总体分布的形式，但不知其参数的情况，提出参数的假设，并根据样本进行检验．

即参数的假设检验

（2）在总体的分布函数完全未知的情况下，提出总体服从某个已知分布的假设，并根据样本进行检验．

即非参数的假设检验，也即分布的假设检验

例　设根据以往经验，某种清漆的干燥时间（以小时计）服
从 $N(6.0,0.36)$，现抽取某批次的 9 个样品：

6.9　6.7　5.8　7.0　6.8　5.2　7.1　5.6　6.5，

问该批次清漆的平均干燥时间是否与以往有显著差异?

例　一种摄影药品被其制造商声称其贮藏寿命是均值 180 天、标准差不多于 10 天的正态分布。某位使用者担心标准差可能超过 10 天。他随机选取 12 个样品并测试，得到样本标准差为 14 天。根据样本有充分证据证明标准差大于 10 天吗?

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、这些数据提供充分证据拒绝该理论吗?

例 1 设根据以往经验，某种清漆的干燥时间(以小时计)

服从 N(6.0,0.36)，现抽取某批次的 9 个样品： 6.9

6.7 5.8 7.0 6.8 5.2 7.1 5.6 6.5 ，问该批次清漆的平均干燥时间是否与以往有显著差异？

(分析)这里我们研究的总体是什么？是服从$N(6.0,0.36)$的吗？

对哪个参数作检验？

这个批次清漆的干燥时间构成的总体方差可设$\sigma^2 = 0.36$

而其均值是要求我们检验的！

经计算，现抽取的9个数据的平均值$\bar{x} = 6.4$小时，

现在的问题是，我们能否认为 "$6.4 \approx 6.0 = \mu_0$" ？

即，接受以下哪个假设？

原假设 $H_0 : \mu = \mu_0 = 6.0$， 备择假设 $H_1 : \mu \neq \mu_0 = 6.0$

原假设 $H_0: \mu = \mu_0 = 6.0$，　备择假设 $H_1: \mu \neq \mu_0 = 6.0$

因为 $\overline{X}$ 是 μ 的矩估计和最大似然估计,而且还是无偏估计，

所以，从总体中抽取的$\overline{X}$在一定程度上反映了总体参数μ 。

反证法

若 H_0 为真,则 $\overline{X}$ 与 μ_0 相差不应太大,否则有理由拒绝H_0 。
所以拒绝 H_0 的形式为 $\left|\overline{X} - \mu_0\right| \geq C.$

检验规则：

当$|\overline{X} - \mu_0| \geq C$时，拒绝原假设$H_0$;

当$|\overline{X} - \mu_0| < C$时，接受原假设$H_0$,

其中C是待定的常数.

用于判断原假设H_0是否成立的统计量$T=T(X_1,\cdots,X_n)$称为对应假设问题的**检验统计量**，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**.

上述例子中，可取检验统计量为$\overline{X}$ (或$\overline{X}-\mu_0$),

拒绝域为 $W=\left\{(X_1,\cdots,X_n):|\overline{X}-\mu_0|\geq C\right\}$

现在的问题是，如何确定临界值C?

也要注意，即使确定了临界值C，但也不能保证判断一定正确，都有可能犯错误。

两类错误

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。

决策	原假设 H_0	
	真的	假的
接受 H_0	正确决策	第二类错误
拒绝 H_0	第一类错误	正确决策

第一类错误：原假设 H_0 成立，　　作出拒绝原假设的决策；
弃真

第二类错误：原假设 H_0 不成立，作出接受原假设的决策。
存伪

犯第I类错误的概率：

$$P\{\text{第}I\text{类错误}\}=P\{\text{拒绝}H_0 \mid H_0\text{为真}\}$$

$$\overset{\text{记为}}{=} P_{H_0\text{为真}}\{\text{拒绝}H_0\}$$

犯第II类错误的概率：

$$P\{\text{第}II\text{类错误}\}=P\{\text{接受}H_0 \mid H_0\text{为假}\}$$

$$\overset{\text{记为}}{=} P_{H_0\text{为假}}\{\text{接受}H_0\}$$

犯第I类错误的概率计算

例1中，犯第I类错误的概率

$$\alpha(C)=P\{拒绝H_0 | H_0为真\}$$

$$=P\{|\bar{X}-\mu_0|\ge C \mid \mu=\mu_0\} \overset{记为}{=} P_{\mu=\mu_0}\{|\bar{X}-\mu_0|\ge C\}$$

$$=P_{\mu=\mu_0}\{\frac{|\bar{X}-\mu_0|}{\sigma/\sqrt{n}}\ge\frac{C}{\sigma/\sqrt{n}}\}$$

$$=1-P_{\mu=\mu_0}\{\left|\frac{\bar{X}-\mu_0}{\sigma/\sqrt{n}}\right|<\frac{C}{\sigma/\sqrt{n}}\}=2-2\Phi\left(\frac{C}{\sigma/\sqrt{n}}\right)$$

$\because$当H_0成立时，$\mu=\mu_0$，$\frac{\bar{X}-\mu_0}{\sigma/\sqrt{n}}\sim N(0,1)$

犯第II类错误的概率

$\beta(C)=P\{$接受$H_0 | H_0$为假$\}$

$$=P_{\mu\neq\mu_0}\{|\bar{X}-\mu_0|<C\}$$

$$=P_{\mu\neq\mu_0}\{\mu_0-C<\bar{X}<\mu_0+C\}$$

$$=P_{\mu\neq\mu_0}\{\frac{\mu_0-C-\mu}{\sigma/\sqrt{n}}<\frac{\bar{X}-\mu}{\sigma/\sqrt{n}}<\frac{\mu_0+C-\mu}{\sigma/\sqrt{n}}\}$$

$$=\Phi\left\{\frac{\mu_0+C-\mu}{\sigma/\sqrt{n}}\right\}-\Phi\left\{\frac{\mu_0-C-\mu}{\sigma/\sqrt{n}}\right\},\quad \mu\neq\mu_0$$

$\therefore$只有μ已知才能知道犯第II类错误的概率！

$$\alpha(C)=2-2\Phi\left(\frac{C}{\sigma/\sqrt{n}}\right)$$

$$\beta(C)=\Phi\left\{\frac{\mu_0+C-\mu}{\sigma/\sqrt{n}}\right\}-\Phi\left\{\frac{\mu_0-C-\mu}{\sigma/\sqrt{n}}\right\}$$

显然，犯第I类错误的概率$\alpha(C)$关于C是单调减函数，而犯第II类错误的概率$\beta(C)$关于C是单调增函数.

在样本容量n固定时，不可能找到界值C，同时使得$\alpha(C)$和$\beta(C)$都很小.即，犯两类错误的概率相互制约！

实际上，需要同时使得$\alpha(C)$和$\beta(C)$都在要求范围内时，可以增加样本容量！

奈曼-皮尔逊($Neyman$ - $Pearson$)原则：

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

其中的常数α称为显著水平,

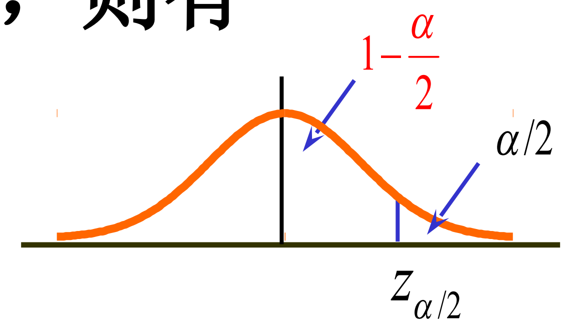
是允许犯第一类错误的最大概率值,

常取$\alpha=0.01,\quad 0.05,\quad 0.1$等.

由奈曼-皮尔逊原则确定C

在例1中，若取显著水平为α，则有

$$\alpha(C) = 2 - 2\Phi\left(\frac{C}{\sigma/\sqrt{n}}\right) \le \alpha,$$



$$\Phi\left(\frac{C}{\sigma/\sqrt{n}}\right) \ge 1-\frac{\alpha}{2}, \qquad \frac{C}{\sigma/\sqrt{n}} \ge z_{\alpha/2},$$

$$C \ge z_{\alpha/2}\sigma/\sqrt{n}\ .$$

由于犯第II类错误的概率$\beta(C)$关于C单调增函数，根据$Neyman-Pearson$原则，应取以上的最小值：$\qquad C = z_{\alpha/2}\sigma/\sqrt{n}$

因此,假设检验“ $H_0: \mu = \mu_0$, $H_1: \mu \neq \mu_0$” 中

H_0的拒绝域为：$W = \left\{ |\bar{X} - \mu_0| \geq z_{\alpha/2}\sigma / \sqrt{n} \right\}$

或写为：$W = \left\{ \left| \dfrac{\bar{X} - \mu_0}{\sigma / \sqrt{n}} \right| \geq z_{\alpha/2} \right\}$

一般简单写为：$\left| \dfrac{\bar{X} - \mu_0}{\sigma / \sqrt{n}} \right| \geq z_{\alpha/2}$

虽然这只是一个样本，但如果一个样本使得上述不等式成立，则可根据实际推断原理（小概率事件几乎在一次试验中不发生），就可以拒绝H_0,否则无理由拒绝H_0。

在例1中，$\mu_0 = 6,\ \sigma = 0.6,\ n = 9,\ \bar{x} = 6.4$，

若取$\alpha = 0.05$, 则 $z_{\alpha/2} = z_{0.025} = 1.96$

$$C = z_{\alpha/2}\sigma / \sqrt{n} = 1.96*0.6*/\sqrt{9} = 0.392$$

有 $\left|\bar{x} - 6\right| = 0.4 > 0.392$,样本落入拒绝域.

也有$\left|\dfrac{\bar{x} - \mu_0}{\sigma / \sqrt{n}}\right| = \left|\dfrac{6.4 - 6}{0.6 / \sqrt{9}}\right| = 2 > z_{0.025}$,样本落入拒绝域.

所以可拒绝H_0，也即我们有**95%**的把握认为该批次清漆的平均干燥时间与以往有显著差异！

根据上述检验规则，犯第I类错误的概率

$$\alpha(C)=\alpha(0.392)=2-2\Phi\left(\frac{0.392}{0.6/\sqrt{9}}\right)=0.05=\alpha$$

犯第II类错误的概率

$$\beta(0.392)=\Phi\left\{\frac{6.0+0.392-\mu}{0.6/\sqrt{9}}\right\}-\Phi\left\{\frac{6.0-0.392-\mu}{0.6/\sqrt{9}}\right\}$$

$$=\Phi\left\{\frac{6.392-\mu}{0.2}\right\}-\Phi\left\{\frac{5.608-\mu}{0.2}\right\},\qquad \mu\neq 6.0$$

例如，当$\mu=6.5$时，犯第II类错误的概率

$$\beta(0.392)=\Phi(-0.054)-\Phi(-0.446)\approx 0.15$$

假设类型：

原假设(零假设）H_0，备择假设（对立假设）H_1

关于总体参数θ的假设有三种：

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

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

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

写法约定：(1) 等号总是在H_0 (2)总体参数写在前

左边假设问题

反证法

$H_0: \mu \ge \mu_0, H_1: \mu < \mu_0$， 其中$\mu_0$是已知的常数

以$\bar{X}$作为μ的参考,若H_0为真，$\bar{X}$比μ_0大些，但是，由于样本的随机波动性，也有可能导致“$\bar{X} < \mu_0$”发生，但如果“$\bar{X} - \mu_0 \le k$”时，则认为H_0为假。因为同样含有“$\bar{X} - \mu_0$”,所以,

检验统计量仍取为 $Z = \dfrac{\bar{X} - \mu_0}{\sigma/\sqrt{n}}$ (与双边检验相同)

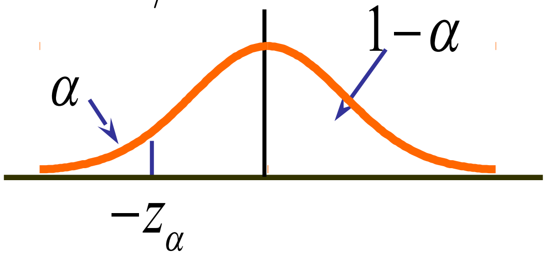
拒绝域形式可表示为 $Z = \dfrac{\bar{X} - \mu_0}{\sigma/\sqrt{n}} \le C$

左边检验“ $H_0: \mu \ge \mu_0, H_1: \mu < \mu_0$”，$H_0$拒绝域的确定

$$\alpha(\mu, C) = P\{拒绝H_0 | H_0是真的\} = P_{\mu\ge\mu_0}\{\frac{\bar{X}-\mu_0}{\sigma/\sqrt{n}} \le C\}$$

$$= P_{\mu\ge\mu_0}\{\frac{\bar{X}-\mu}{\sigma/\sqrt{n}} \le C - \frac{\mu-\mu_0}{\sigma/\sqrt{n}}\} = \Phi_{\mu\ge\mu_0}(C - \frac{\mu-\mu_0}{\sigma/\sqrt{n}})$$

$$\le \Phi(C - \frac{\mu_0-\mu_0}{\sigma/\sqrt{n}}) = \Phi(C) \overset{设}{\le} \alpha$$



以上也可写为$\sup\limits_{\mu\ge\mu_0} \alpha(\mu, C) = \alpha(\mu_0, C) = \Phi(C) \overset{设}{\le} \alpha$

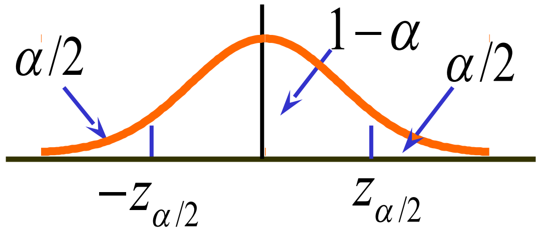
$$\Rightarrow C = -z_\alpha$$

$\therefore$左边检验的拒绝域为 $Z = \dfrac{\bar{X}-\mu_0}{\sigma/\sqrt{n}} \le -z_\alpha$

三种假设检验的拒绝域

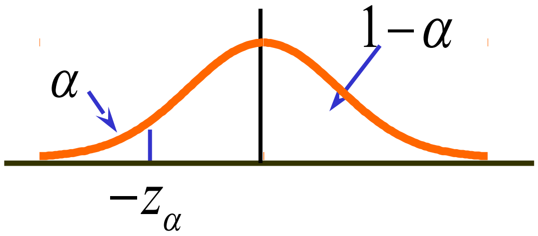
双边假设问题 $H_0: \mu = \mu_0, H_1: \mu \neq \mu_0$

拒绝域为$|Z| = \left|\dfrac{\bar{X} - \mu_0}{\sigma/\sqrt{n}}\right| \geq z_{\alpha/2}$



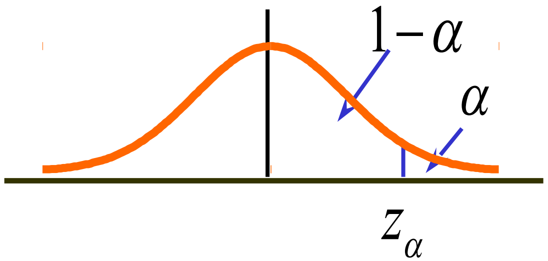
左边假设问题 $H_0: \mu \geq \mu_0, H_1: \mu < \mu_0$

拒绝域为$Z = \dfrac{\bar{X} - \mu_0}{\sigma/\sqrt{n}} \leq -z_\alpha$



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

拒绝域为$Z = \dfrac{\bar{X} - \mu_0}{\sigma/\sqrt{n}} \geq z_\alpha$



$P_$值定义及计算（$P215$）

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

计算： 例1中的检验 $H_0: \mu = \mu_0 = 6.0$, $H_1: \mu \neq \mu_0$

H_0的拒绝域为: $|Z| = \left|\dfrac{\overline{X} - \mu_0}{\sigma/\sqrt{n}}\right| \geq z_{\alpha/2}$,

$$|z_0| = \left|\frac{\overline{x} - \mu_0}{\sigma/\sqrt{n}}\right| = \left|\frac{6.4 - 6.0}{0.6/\sqrt{9}}\right| = 2,$$

$$P\_ = P_{H_0}\left(|Z| \geq |z_0|\right) = P_{H_0}\left(|Z| \geq 2\right) = 2 - 2\Phi(2) = 0.046$$

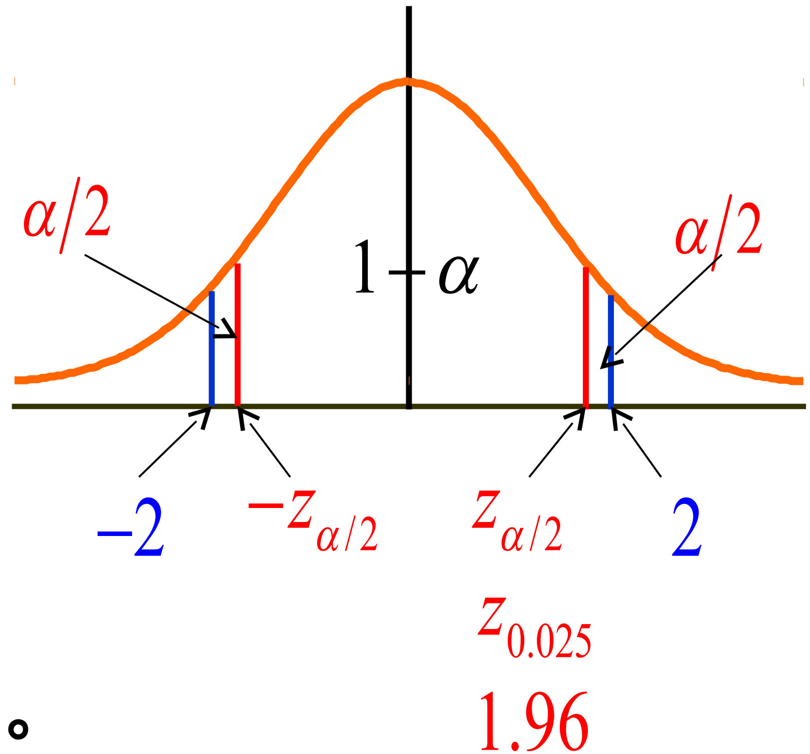
说明： 统计量Z的值绝对值$|z_0|$取代拒绝域中的$z_{\alpha/2}$。

用P_值判断实例

例1中，若取显著性水平α=0.05, 则由

$$P\_ = P_{H_0}\left(|Z|\geq 2\right) = 0.046 \quad < \quad \alpha = 0.05$$

可知，样本数据必定落在H_0的拒绝域内。所以应拒绝原假设！



事实上，由图也可看出样本数据落在拒绝域内。

$P_$值与统计显著性

在实际运用中，通过$P_$值来衡量拒绝H_0的理由是否充分。$P_$值较小说明观察的结果在一次试验中发生的可能性较小，值越小，拒绝H_0的理由越充分；值越大，越没有足够的理由拒绝H_0。

一般，$P_$值与显著性水平α作比较，

若$P_ \le \alpha$，则拒绝原假设；

若$P_ > \alpha$，则接受原假设.

拒绝原假设，称在水平α下统计显著；

接受原假设，称在水平α下统计不显著。

左边Z假设检验的$P_$值计算

参照左边检验的拒绝域 $Z=\dfrac{\bar{X}-\mu_0}{\sigma/\sqrt{n}}\le -z_\alpha$

对给定的样本观察值$x_1,\cdots,x_n$,

计算检验统计量值 $z_0=\dfrac{\bar{x}-\mu_0}{\sigma/\sqrt{n}}$.

$$P\_=\sup_{\mu\ge\mu_0}P\left\{Z\le z_0\right\}=\sup_{\mu\ge\mu_0}P\left\{\frac{\bar{X}-\mu_0}{\sigma/\sqrt{n}}\le z_0\right\}$$

$$=P\left\{Z\le z_0\mid\mu=\mu_0\right\}=\Phi(z_0)$$

μ代替μ_0后概率变大

简写：$P\_=P(Z\le z_0)=\Phi(z_0)$

右边Z假设检验的$P_$值计算

对给定的样本观察值$x_1,\cdots,x_n$， $z_0=\dfrac{\bar{x}-\mu_0}{\sigma/\sqrt{n}}$.

（拒绝域为$Z=\dfrac{\bar{X}-\mu_0}{\sigma/\sqrt{n}}\geq z_\alpha$）

$$P\_=\sup_{\mu\leq\mu_0}P_{H_0}\{Z\geq z_0\}=\sup_{\mu\leq\mu_0}P_{H_0}\left\{\frac{\bar{X}-\mu_0}{\sigma/\sqrt{n}}\geq z_0\right\}$$

$$=P\{Z\geq z_0\mid\mu=\mu_0\}$$

$$=1-\Phi(z_0).$$

μ代替μ_0后概率变大

简写：$P\_=P(Z\geq z_0)=1-\Phi(z_0)$

参数的假设检验问题处理步骤

- 1. 根据实际问题的要求，提出原假设 H_0 和备择假设 H_1；特别注意 H_0 与 H_1 的不对等性。
- 2. 根据已知条件选取检验统计量（方法与区间估计中枢轴量的选取相同），画出统计量密度函数草图；
- 3. 按照"在原假设 H_0 成立时，拒绝原假设的概率不大于显著性水平α"这一原则，画出统计量分布的分位数图，（左边检验左边留α，右边检验右边留α，两边检验两边各留α/2）确定 H_0 拒绝域；
- 4．查分位数表、用样本观测值数据代入公式进行计算；根据样本数据是否落在 H_0 拒绝域内，作出拒绝原假设还是接受原假设的决策。

（$3'$）按"3."确定拒绝域，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 与 P_大小，作出判断.

H_0和H_1的不对等性

在进行显著性检验时，犯第Ⅰ类错误的概率是由我们控制的．α 取得小，则概率 $P\{$当 H_0 为真拒绝 $H_0\}$就小，这保证了当 H_0 为真时错误地拒绝 H_0 的可能性很小．这意味着 H_0 是受到保护的，也表明 H_0、H_1 的地位不是对等的．于是，在一对对立假设中，选哪一个作为 H_0 需要小心．例如，考虑某种药品是否为真，这里可能犯两种错误：(1)将假药误作为真药，则冒着伤害病人的健康甚至生命的风险；(2)将真药误作为假药，则冒着造成经济损失的风险．显然，犯错误(1)比犯错误(2)的后果严重，因此，我们选取"H_0：药品为假，H_1：药品为真"，即是使得犯第Ⅰ类错误"当药品为假时错判药品为真"的概率$\leqslant\alpha$．就是说，选择 H_0、H_1 使得两类错误中后果严重的错误成为第Ⅰ类错误．这是选择 H_0、H_1 的一个原则．

如果在两类错误中，没有一类错误的后果严重更需要避免时，常常取 H_0 为维持现状，即

H_0和H_1的不对等性

取 H_0 为“无效益”、“无改进”、“无价值”等等．例如，取

$$H_0\text{:新技术未提高效益}, \quad H_1\text{:新技术提高效益}.$$

实际上，我们感兴趣的是 H_1“提高效益”，但对采用新技术应持慎重态度．选取 H_0 为“新技术未提高效益”，一旦 H_0 被拒绝了，表示有较强的理由去采用新技术．

在实际问题中，情况比较复杂，如何选取 H_0、H_1 只能在实践中积累经验，根据实际情况去判断了．

一般，在有关参数的假设检验中，备择假设是我们根据样本资料希望得到支持的假设。

或：题目中问的问题写到 H_1 中

正态总体下常用的检验统计量

一、单个总体$N(\mu,\sigma^2)$下常用检验统计量

（1）σ^2已知,对μ的检验： $\dfrac{\bar{X}-\mu_0}{\sigma/\sqrt{n}} \sim N(0,1)$

（2）σ^2未知,对μ的检验： $\dfrac{\bar{X}-\mu_0}{S/\sqrt{n}} \sim t(n-1)$

（3）μ未知,对σ^2的检验： $\dfrac{(n-1)S^2}{\sigma_0^2} \sim \chi^2(n-1)$

（4）μ已知,对σ^2的检验： $\sum_{i=1}^{n}\left(\dfrac{X_i-\mu}{\sigma_0}\right)^2 \sim \chi^2(n)$

正态总体下常用的检验统计量

二、双总体$N(\mu_1,\sigma_1^2),N(\mu_2,\sigma_2^2)$下常用检验统计量

(1) σ_1^2,σ_2^2已知,对$H_0:\mu_1-\mu_2=\delta$的检验:$\dfrac{(\bar{X}-\bar{Y})-\delta}{\sqrt{\dfrac{\sigma_1^2}{n_1}+\dfrac{\sigma_2^2}{n_2}}}\sim N(0,1)$

(2) $\sigma_1^2=\sigma_2^2$未知,对$H_0:\mu_1-\mu_2=\delta$的检验:$\dfrac{(\bar{X}-\bar{Y})-\delta}{S_w\sqrt{\dfrac{1}{n_1}+\dfrac{1}{n_2}}}\sim t(n_1+n_2-2)$

(3) μ_1,μ_2未知,对$H_0:\sigma_1^2=\sigma_2^2$的检验:$\dfrac{S_1^2}{S_2^2}\sim F(n_1-1,n_2-1)$

(4) μ_1,μ_2已知,对$H_0:\sigma_1^2=\sigma_2^2$的检验:$\dfrac{\sum_{i=1}^{n_1}(X_i-\mu_1)^2\Big/n_1}{\sum_{i=1}^{n_2}(Y_i-\mu_2)^2\Big/n_2}\sim F\left(n_1,n_2\right)$

§8.2 单个正态总体参数的假设检验

$X_1, X_2, \cdots, X_n$来自$N\left(\mu, \sigma^2\right)$, $\bar{X}$和S^2分别为样本均值和方差,显著性水平为 α

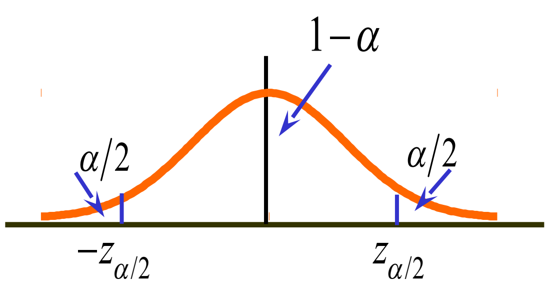
(一) 单个正态总体$N\left(\mu, \sigma^2\right)$均值$\mu$的检验

1. σ^2已知时的双边检验 $H_0: \mu = \mu_0, H_1: \mu \neq \mu_0$

取检验统计量 $Z = \dfrac{\bar{X} - \mu_0}{\sigma / \sqrt{n}}$ $\overset{\text{在}H_0\text{为真时}}{\sim}$ $N\left(0,1\right)$

H_0的拒绝域为:

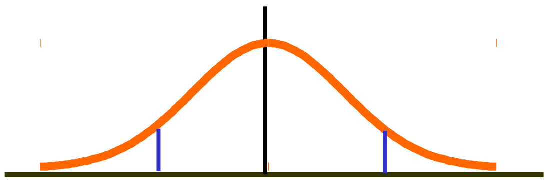
$$\boxed{|Z| = \left|\frac{\bar{X} - \mu_0}{\sigma / \sqrt{n}}\right| \geq z_{\alpha/2}}$$



称此检验为“ Z检验法”

利用$P_$值判断H_0真伪

参照H_0的拒绝域 $|Z|=\left|\dfrac{\bar{X}-\mu_0}{\sigma/\sqrt{n}}\right|\geq z_{\alpha/2}$,



对给定的样本观察值$x_1,\cdots,x_n$,

记检验统计量Z的取值为 $z_0=\dfrac{\bar{x}-\mu_0}{\sigma/\sqrt{n}}$,

$$P\_=P_{H_0}\{|Z|\geq|z_0|\}=2P_{H_0}\{Z\geq|z_0|\}=2(1-\Phi(|z_0|))$$

(即$|z_0|$代替了拒绝域式中的$z_{\alpha/2}$)

判断：当$P_$小于显著水平α时，拒绝原假设，

否则，接受原假设.

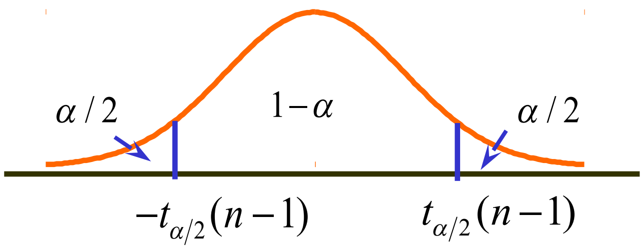
2. σ^2未知时的双边检验　$H_0:\mu=\mu_0, H_1:\mu\neq\mu_0$

由于σ^2未知，故不能用$\dfrac{\bar{X}-\mu_0}{\sigma/\sqrt{n}}$来确定拒绝域了。

取检验统计量：$t=\dfrac{\bar{X}-\mu_0}{S/\sqrt{n}} \overset{\text{在}H_0\text{为真时}}{\sim} t(n-1)$

H_0的拒绝域为：

$$\boxed{|t|=\frac{|\bar{X}-\mu_0|}{S/\sqrt{n}}\geq t_{\alpha/2}(n-1)}$$



称此检验为t检验法

双边t假设检验的$P_$值计算

对给定的样本观察值$x_1,\cdots,x_n$，记检验统计量t的取值为

$t_0=\dfrac{\bar{x}-\mu_0}{s/\sqrt{n}}$，则有（拒绝域为：$|t|\ge t_{\alpha/2}(n-1)$ ）

$$P\_=P_{H_0}(|t|\ge|t_0|)=P_{H_0}(\{t\ge|t_0|\}\cup\{t\le-|t_0|\})$$

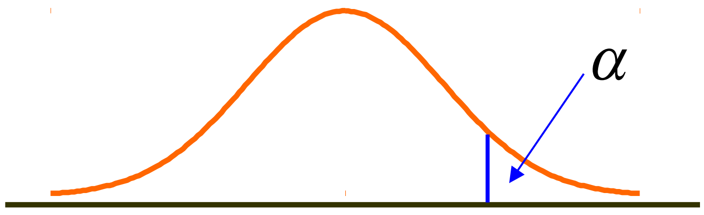
$$=2P(t\ge|t_0|)\overset{记}{=}2P(t(n-1)\ge|t_0|).$$

当$P_\le\alpha$时，拒绝原假设，

当$P_>\alpha$时，接受原假设.

2. σ^2未知时的右边检验 $H_0: \mu \le \mu_0, H_1: \mu > \mu_0$

取检验统计量：$t = \dfrac{\bar{X} - \mu_0}{S/\sqrt{n}}$



H_0的拒绝域为：

$$\boxed{t = \frac{\bar{X} - \mu_0}{S/\sqrt{n}} \ge t_\alpha(n-1)}$$

$$P\_ = \sup_{\mu \le \mu_0} \{ t \ge t_0 \} = P\{ t(n-1) \ge t_0 \}.$$

思考题：左边检验

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

H_0的拒绝域是什么？

答：拒绝域为：

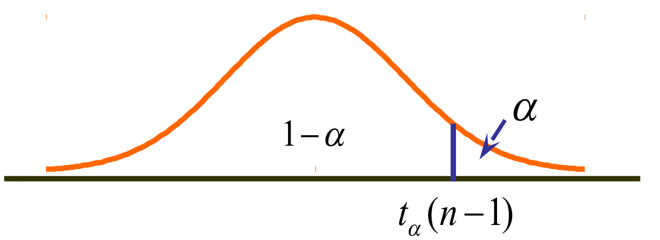
$$\frac{\bar{X} - \mu_0}{S/\sqrt{n}} \le -t_\alpha(n-1)$$

$$P\_ = \sup_{\mu \ge \mu_0} \{ t \le t_0 \}$$

$$= P\{ t(n-1) \le t_0 \}.$$

例2：某种元件的寿命X（以小时记）服从正态分布$N(\mu,\sigma^2)$, μ,σ^2均未知。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 ，问可否认为元件的平均寿命大于225小时？(取$\alpha=0.05$)

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



取检验统计量：$t=\dfrac{\bar{X}-\mu_0}{S/\sqrt{n}}$

H_0的拒绝域为：　$t\ge t_\alpha(n-1).$

$$n=16,\ t_{0.05}(15)=1.7531,\ \bar{x}=241.5,\ s=98.7259$$

计算得：$t=\dfrac{\bar{X}-\mu_0}{S/\sqrt{n}}=0.6685<1.7531=t_{0.05}(15).$

t 没有落在拒绝域内，故接受原假设，

即认为元件的平均寿命不大于 **225** 小时。

*$Excel$计算$P_$值

由$Excel$可计算$P_$值为

$$P_{-} = P_{H_0}\left\{t \geq t_0\right\} = P\left\{t(15) \geq 0.6685\right\}$$

=T.DIST.RT(0.6685,15)

$$\approx 0.257 > 0.05$$

因此接受原假设，

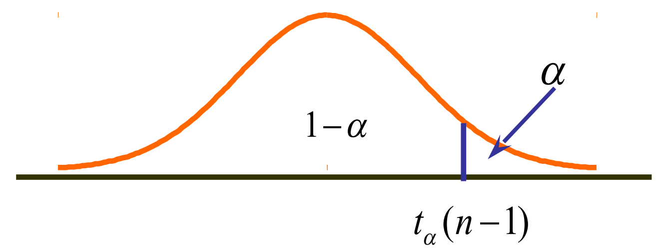
即认为元件的平均寿命不大于**225**小时。

判断结果与前面一致！

例3 要求某元件的平均寿命不得低于1000小时，现从一批这种元件中随机抽取25件，测得其平均寿命为980小时，标准差为100小时。已知这批元件的寿命服从正态分布。试在显著性水平0.05下确定这批元件是否合格？

解： $H_0: \mu \le \mu_0 = 1000$，　$H_1: \mu > 1000$

取检验统计量： $t = \dfrac{\bar{X} - \mu_0}{S/\sqrt{n}}$



H_0的拒绝域为： $t \ge t_\alpha(n-1).$

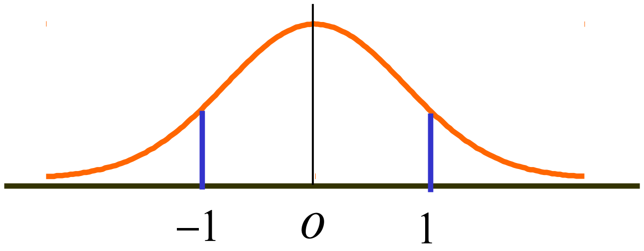
$$n = 25, t_{0.05}(24) = 1.7109. \quad \bar{x} = 980, s = 100$$

计算得： $t = \dfrac{\bar{X} - \mu_0}{S/\sqrt{n}} = -1 < 1.7109 = t_{0.05}(24).$

t 不落在拒绝域内，故接受原假设 H_0，即认为这批元件的平均寿命小于 **1000** 小时，不合格。

*$Excel$计算$P_$值

H_0的拒绝域为：　$t \ge t_{\alpha}(n-1).$



$P_$值为

$$P\_ = P_{H_0}\left\{t \ge t_0\right\} = P\left\{t(24) \ge -1\right\}$$

$$= P\left\{t(24) \le 1\right\} \overset{Excel}{=} \textbf{T.DIST(1,24,1)} \approx \mathbf{0.84}$$

$$= 1 - P\left\{t(24) > 1\right\}$$

$$\overset{Excel}{=} 1 - \textbf{T.DIST.RT(1,24)}$$

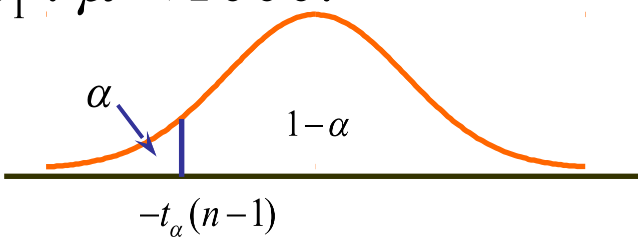
$$\approx 1 - 0.16 = 0.84 > 0.05$$

因此接受原假设，判断结果与前面一致！

例3′ 要求某元件的平均寿命不得低于1000小时，现从一批这种元件中随机抽取25件，测得其平均寿命为980小时，标准差为100小时。已知这批元件的寿命服从正态分布。试在显著性水平0.05下确定这批元件是否合格？

解：如果检验设为 $H_0: \mu \geq \mu_0 = 1000$，$H_1: \mu < 1000.$

取检验统计量：$t = \dfrac{\bar{X} - \mu_0}{S/\sqrt{n}}$



H_0的拒绝域为：$t \leq -t_\alpha(n-1)$

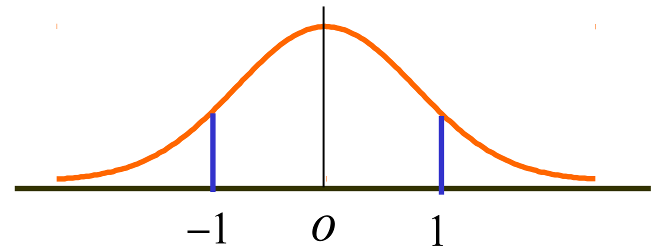
$$n = 25, t_{0.05}(24) = 1.7109. \quad \bar{x} = 980, s = 100$$

计算得：$t = \dfrac{\bar{X} - \mu_0}{S/\sqrt{n}} = -1 > -1.7109 = -t_{0.05}(24).$

t 不落在拒绝域内，故接受原假设 **H**$_0$，即认为这批元件的平均寿命不小于 **1000** 小时，合格。

*$Excel$计算$P_$值

H_0的拒绝域为：$t \le -t_\alpha(n-1)$



$P_$值为

$$P\_ = P_{H_0}\{t \le t_0\} = P\{t(24) \le -1\} \overset{Excel}{=} \textbf{T.DIST}(-1,24,1)$$

$$= P\{t(24) > 1\}$$

$$\overset{Excel}{=} \textbf{T.DIST.RT}(1,24)$$

$$\approx 0.16 > 0.05$$

因此接受原假设，判断结果与前面一致！

（二）成对数据的t检验

成对数据在区间估计中已作过介绍.

成对样本设为 $(X_1, Y_1), \cdots, (X_n, Y_n)$,

差值为 $D_i = X_i - Y_i, \quad i = 1, \cdots, n.$

可以看成来自正态总体$N(\mu_d, \sigma_d^2)$的样本。

为比较两总体均值是否有显著差异，可考虑假设问题

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

转化为单个正态总体均值的假设检验。

记 $\bar{D}=\frac{1}{n}\sum_{i=1}^{n}D_i$, $S_d^2=\frac{1}{n-1}\sum_{i=1}^{n}(D_i-\bar{D})^2$

则检验统计量为 $t=\frac{\bar{D}}{S_d/\sqrt{n}}$,

检验的拒绝域为 $|t|\geq t_{\alpha/2}(n-1)$

$P_$值为

$$P\_=P_{H_0}\left\{|t|\geq|t_0|\right\}=2P\left\{t(n-1)\geq|t_0|\right\}$$

例 4　为了试验两种不同谷物种子的优劣，选取了十块土质不同的土地，并将每块土地分为面积相同的两部分，分别种植这两种种子。设在每块土地的两部分人工管理等条件完全一样。下面给出各块土地上的产量。

土地	1	2	3	4	5	6	7	8	9	10
种子 A(x_i)	23	35	29	42	39	29	37	34	35	28
种子 B(y_i)	26	39	35	40	38	24	36	27	41	27
$d_i=x_i-y_i$	−3	−4	−6	2	1	5	1	7	−6	1

问：以这两种种子种植的谷物产量是否有显著的差异？（取显著性水平为 0.05）

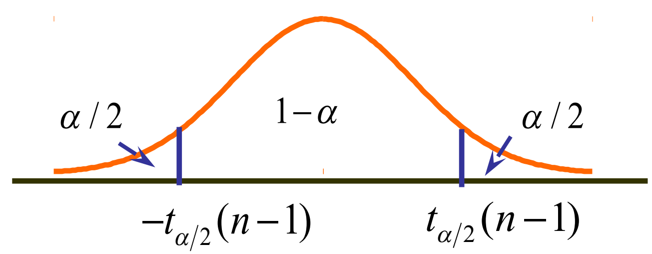
分析：不能认为某一种子数据是正态的！本题中每对数据的差异仅是由这两种种子的差异造成的，将每对数据作差就能排除其它因素的影响，从而能比较种子产量的差异。一般，设有n对独立的观测结果：$(X_1,Y_1),(X_2,Y_2),\dots,(X_n,Y_n)$，令 $D_1=X_1-Y_1, D_2=X_2-Y_2,\dots,D_n=X_n-Y_n$，

则$D_1,D_2,\dots,D_n$相互独立，服从同一正态分布，设为$N(\mu_d,\sigma_d^2)$.

解：检验假设 $H_0:\mu_d=0,\ H_1:\mu_d\neq 0$

分别将$D_1,D_2,\dots,D_n$的样本均值和样本方差的观测值记为$\bar{d}$，S_d^2，

取统计量$t=\dfrac{\bar{D}}{S_d/\sqrt{n}}\overset{H_0}{\sim}t(n-1)$



H_0的拒绝域为：$|t|\geq t_{\alpha/2}(n-1)$，

$n=10$，$t_{0.025}(9)=2.2622$，$\bar{d}=-0.2$，$S_d=4.442$，

计算得：$\dfrac{|\bar{d}|}{S_d/\sqrt{n}}=0.142<2.2622$

接受原假设H_0，认为两种种子的产量没有显著差异。

H_0的拒绝域为：$|t| \geq t_{\alpha/2}(n-1)$，

$P_$值为

$$P\_ = P_{H_0}\{|t| \geq |t_0|\} = 2P\{t(9) \geq 0.142\}$$

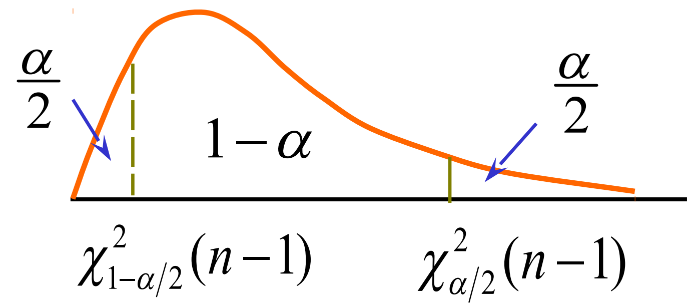
$$\overset{Excel}{=} 2*\textbf{T.DIST.RT(0.142,9)}=2*0.445$$

$$> 0.05$$

因此接受原假设，判断结果与前面一致！

（三）单个正态总体方差σ^2的假设检验

$$H_0:\sigma^2=\sigma_0^2,\ H_1:\sigma^2\neq\sigma_0^2$$ 设μ未知



取检验统计量：

$$\chi^2=\frac{(n-1)S^2}{\sigma_0^2}\overset{H_0\text{为真}}{\sim}\chi^2(n-1)$$

H_0的拒绝域为：

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

称此检验为χ^2检验法。

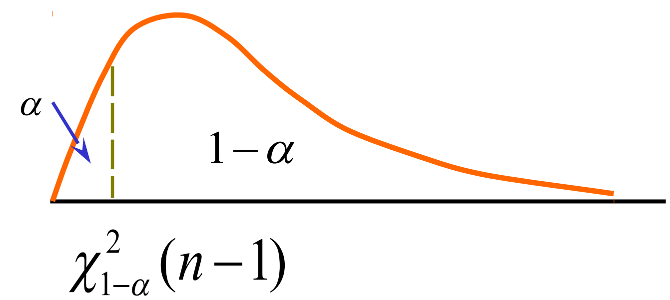
记$\chi_0^2=(n-1)S^2/\sigma_0^2$，分布不对称的双边检验$P$值：

$$P\_=2\min\left\{P[\chi^2(n-1)\le\chi_0^2],P[\chi^2(n-1)\ge\chi_0^2]\right\}$$

即，若记 $p_0=P\{\chi^2(n-1)\le\chi_0^2\}$，则$P_=2\min\{\ p_0,1-p_0\}$

$$H_0:\sigma^2 \ge \sigma_0^2,\quad H_1:\sigma^2 < \sigma_0^2$$ 设μ未知

取检验统计量 $\chi^2=\dfrac{(n-1)S^2}{\sigma_0^2}$

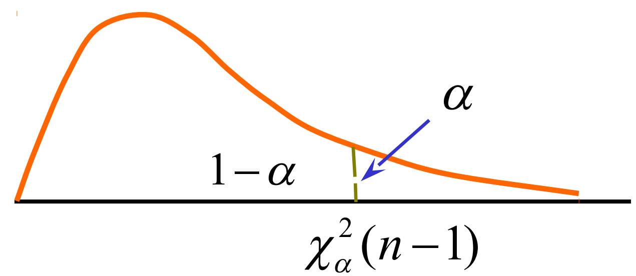


H_0**的拒绝域为:** $\boxed{\dfrac{(n-1)S^2}{\sigma_0^2} \le \chi^2_{1-\alpha}(n-1)}$

$$P\_ = P(\chi^2(n-1) \le \chi_0^2) = P(\chi^2(n-1) \le \frac{(n-1)s^2}{\sigma_0^2})$$

$$H_0:\sigma^2\leq\sigma_0^2,\quad H_1:\sigma^2>\sigma_0^2$$ 设μ未知

取检验统计量 $\chi^2=\dfrac{(n-1)S^2}{\sigma_0^2}$



H_0的拒绝域为：$\boxed{\dfrac{(n-1)S^2}{\sigma_0^2}\geq\chi_\alpha^2(n-1)}$

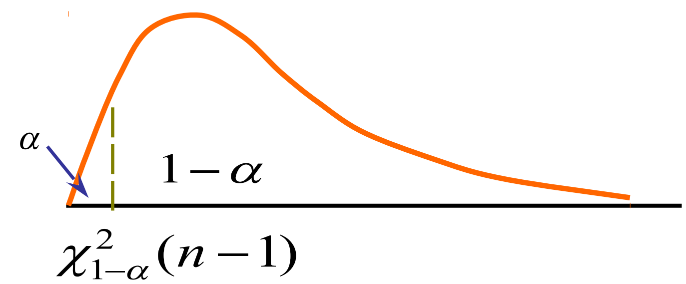
$$P\_=P\left(\chi^2(n-1)\geq\chi_0^2\right)=P\left(\chi^2(n-1)\geq\frac{(n-1)s^2}{\sigma_0^2}\right)$$

例6: 设某苹果新品种其重量是正态分布,这种苹果除了口感好和颜色鲜艳以外,另一个重要特征是单个重量差异不大（对照品种的方差$\sigma^2=7$)。为了评估新苹果,随机挑选了25个测试重量(单位：克),其样本方差为$S^2=4.25.$

在$\alpha=0.05$下检验新品种是否比对照品种方差小?

解：$H_0:\sigma^2\geq\sigma_0^2=7,\quad H_1:\sigma^2<\sigma_0^2=7$

取检验统计量：$\chi^2=\dfrac{(n-1)S^2}{\sigma_0^2}$



H_0的拒绝域为：$\chi^2=\dfrac{(n-1)S^2}{\sigma_0^2}\leq\chi^2_{1-\alpha}(n-1)$

查表得：$\chi^2_{0.95}(24)=13.848,$

计算得：$\dfrac{(25-1)\times4.25}{7}=14.57>13.848$

接受原假设，认为新品种的方差并不比对照组的小。

左边χ^2检验P值计算

H_0的拒绝域为：$\chi^2 \le \chi^2_{1-\alpha}(n-1)$，$\chi_0^2 = 14.57$

$$P\_ = P(\chi^2(24) \le \chi_0^2) = P(\chi^2(24) \le 14.57)$$
$$= 1 - P(\chi^2(24) > 14.57)$$
$$= 1 - \textbf{CHISQ.DIST.RT(14.57,24)}$$
$$= 1 - 0.9327 = 0.0673$$

$$P\_ = 0.0673 > \alpha = 0.05$$

接受原假设，认为新品种的方差并不比对照组的小。

§8.3　两个正态总体参数的假设检验

$X_1, X_2, \cdots, X_{n_1}$来自$N\left(\mu_1, \sigma_1^2\right)$,

$Y_1, Y_2, \cdots, Y_{n_2}$来自$N\left(\mu_2, \sigma_2^2\right)$,

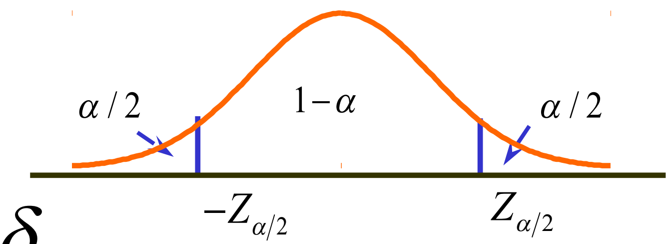
$\bar{X}, \bar{Y}, S_1^2$, S_2^2分别为第一,二个总体的样本均值和方差,显著性水平为 α.

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

(δ为已知常数,通常为0)

（一）比较两个正态总体均值的假设检验

1. 当σ_1^2和σ_2^2已知时



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

取检验统计量：$$Z=\frac{\bar{X}-\bar{Y}-\delta}{\sqrt{\dfrac{\sigma_1^2}{n_1}+\dfrac{\sigma_2^2}{n_2}}}\ \overset{H_0\text{为真}}{\sim}\ N(0,1)$$

则检验拒绝域为：$$|Z|=\frac{\left|\bar{X}-\bar{Y}-\delta\right|}{\sqrt{\dfrac{\sigma_1^2}{n_1}+\dfrac{\sigma_2^2}{n_2}}}\geq z_{\alpha/2}$$ Z检验法

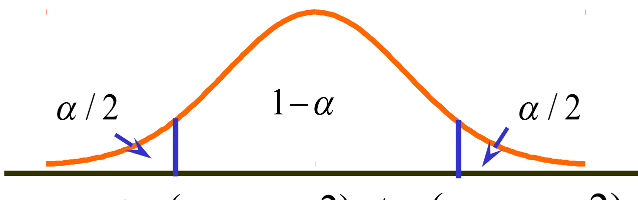
$$P_{-}=P_{H_0}\{|Z|\geq|z_0|\}=2P_{H_0}\{Z\geq|z_0|\}=2(1-\Phi(|z_0|))$$

2. 当$\sigma_1^2=\sigma_2^2=\sigma^2$但未知时

$H_0:\mu_1-\mu_2=\delta, H_1:\mu_1-\mu_2\neq\delta.$ (δ为已知常数)

取检验统计量： $t=\dfrac{\bar{X}-\bar{Y}-\delta}{S_w\sqrt{\dfrac{1}{n_1}+\dfrac{1}{n_2}}}\overset{H_0\text{为真}}{\sim} t(n_1+n_2-2)$

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



H_0**的拒绝域为：** $\boxed{|t|=\dfrac{|\bar{X}-\bar{Y}-\delta|}{S_w\sqrt{\dfrac{1}{n_1}+\dfrac{1}{n_2}}}\geq t_{\alpha/2}(n_1+n_2-2)}$ t **检验法**

H_0的拒绝域为： $|t| = \dfrac{\left|\bar{X} - \bar{Y} - \delta\right|}{S_w\sqrt{\dfrac{1}{n_1} + \dfrac{1}{n_2}}} \geq t_{\alpha/2}\left(n_1 + n_2 - 2\right)$

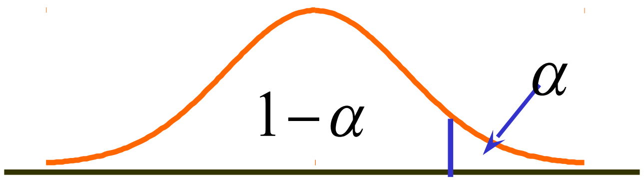
$$t_0 = \frac{\bar{x} - \bar{y} - \delta}{s_w\sqrt{\dfrac{1}{n_1} + \dfrac{1}{n_2}}}$$

$$P_- = P\left\{\,|t| \geq |t_0|\,\right\} = 2P\left\{\,t(n_1 + n_2 - 2) \geq |t_0|\,\right\}$$

———两样本精确t检验

$H_0: \mu_1 - \mu_2 \le \delta, H_1: \mu_1 - \mu_2 > \delta \quad \sigma_1^2 = \sigma_2^2 = \sigma^2, \sigma^2$未知

取检验统计量： $t = \dfrac{\bar{X} - \bar{Y} - \delta}{S_w \sqrt{\dfrac{1}{n_1} + \dfrac{1}{n_2}}}$



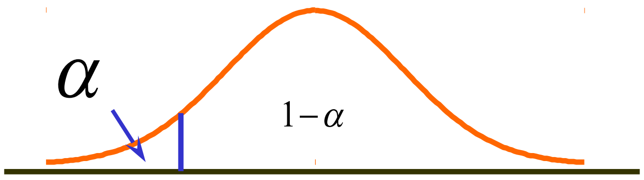
其中 $S_w = \sqrt{\dfrac{(n_1 - 1) S_1^2 + (n_2 - 1) S_2^2}{n_1 + n_2 - 2}}$

当$\mu_1 - \mu_2 = \delta$时， $\dfrac{\bar{X} - \bar{Y} - \delta}{S_w \sqrt{\dfrac{1}{n_1} + \dfrac{1}{n_2}}} \sim t(n_1 + n_2 - 2)$

H_0的拒绝域为： $\boxed{t = \dfrac{\bar{X} - \bar{Y} - \delta}{S_w \sqrt{\dfrac{1}{n_1} + \dfrac{1}{n_2}}} \ge t_\alpha(n_1 + n_2 - 2)}$

$H_0: \mu_1 - \mu_2 \geq \delta, H_1: \mu_1 - \mu_2 < \delta$ $\sigma_1^2 = \sigma_2^2 = \sigma^2, \sigma^2$未知

取检验统计量： $t = \dfrac{\bar{X} - \bar{Y} - \delta}{S_w \sqrt{\dfrac{1}{n_1} + \dfrac{1}{n_2}}}$



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

当$\mu_1 - \mu_2 = \delta$时， $\dfrac{\bar{X} - \bar{Y} - \delta}{S_w \sqrt{\dfrac{1}{n_1} + \dfrac{1}{n_2}}} \sim t(n_1 + n_2 - 2)$

H_0的拒绝域为： $\boxed{t = \dfrac{\bar{X} - \bar{Y} - \delta}{S_w \sqrt{\dfrac{1}{n_1} + \dfrac{1}{n_2}}} \leq -t_\alpha(n_1 + n_2 - 2)}$

*3. 当$\sigma_1^2 \neq \sigma_2^2$且未知时

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

分别以两样本方差S_1^2, S_2^2作为σ_1^2和σ_2^2的无偏估计，取检验统计量为

$$t = \frac{\bar{X} - \bar{Y} - \delta}{\sqrt{\dfrac{S_1^2}{n_1} + \dfrac{S_2^2}{n_2}}}$$

（1）当两样本容量都充分大时，根据大数定律和中心极限定理，当原假设成立时，统计量 t 近似服从标准正态分布 $N(0,1)$.

检验的拒绝域为 $$|t|=\frac{\left|\bar{X}-\bar{Y}-\delta\right|}{\sqrt{\frac{S_1^2}{n_1}+\frac{S_2^2}{n_2}}}\geq z_{\alpha/2}$$

$P\_$ 值为 $P\_=P_{H_0}\left\{|t|\geq|t_0|\right\}=2P\left\{Z\geq|t_0|\right\}$，

其中 $Z\sim N(0,1)$，$t_0=\dfrac{\bar{x}-\bar{y}-\delta}{\sqrt{\frac{s_1^2}{n_1}+\frac{s_2^2}{n_2}}}$.

(2) 对于小样本情形，原假设成立时，统计量t近似服从t分布，自由度为

$$k=\min(n_1-1,n_2-1),$$

或更精确的近似自由度 $k=\dfrac{(S_1^2/n_1+S_2^2/n_2)^2}{\dfrac{(S_1^2/n_1)^2}{n_1-1}+\dfrac{(S_2^2/n_2)^2}{n_2-1}}$

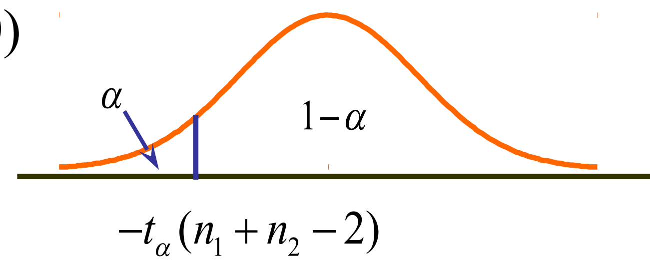
则检验的拒绝域为 $|t|\geq t_{\alpha/2}(k)$

$P_$值为 $P\_=P_{H_0}\left\{|t|\geq|t_0|\right\}=2P\left\{t(k)\geq|t_0|\right\}.$

——————————两样本近似t检验

例5：某厂使用A, B两种不同原料生产同一类型产品。各在一周的产品中取样分析。取用原料A生产的样品220件，测得平均重量为2.46（公斤），样本标准差$s = 0.57$（公斤）。取用原料B生产的样品205件，测得平均重量为2.55（公斤），样本标准差为0.48（公斤）。设两样本独立，来自两个方差相同的独立正态总体。问在水平0.05下能否认为用原料B的生产的产品平均重量较用原料A的大？

解：检验假设 $H_0: \mu_A \geq \mu_B, H_1: \mu_A < \mu_B$ （即$\delta = 0$）



取检验统计量 $t = \dfrac{\bar{X}-\bar{Y}}{S_w\sqrt{\dfrac{1}{n_1}+\dfrac{1}{n_2}}}$

H_0的拒绝域为：$t \leq -t_\alpha(n_1+n_2-2)$

$n_1 = 220,\ \bar{x} = 2.46,\ s_1 = 0.57;\ n_2 = 205,\ \bar{y} = 2.55,\ s_2 = 0.48,\ \alpha = 0.05$

$t_{0.05}(220+205-2) \approx z_{0.05} = 1.645, s_w = 0.535, \sqrt{\dfrac{1}{220}+\dfrac{1}{205}} = 0.097$

$\therefore t = \dfrac{\bar{X}-\bar{Y}}{S_w\sqrt{\dfrac{1}{n_1}+\dfrac{1}{n_2}}} = -1.733 < -1.645$，从而拒绝原假设，即认为…B…比…A大。

H_0的拒绝域为：$t \le -t_\alpha(n_1+n_2-2)$

$$t_0 = \frac{\overline{X}-\overline{Y}}{S_w\sqrt{\frac{1}{n_1}+\frac{1}{n_2}}} = -1.733$$

$$P\_ = P(t(220+205-2) \le t_0) \approx P(Z \le t_0) = \Phi(t_0)$$

$$= \Phi(-1.733) = 1-\Phi(1.733) = 1-0.9582 = 0.0418$$

$P_ < \alpha = 0.05,$从而拒绝原假设。

即在水平**0.05**下可以认为用原料B的生产的产品平均重量较用原料A的大。

（二）比较两个正态总体方差的假设检验

在μ_1, μ_2未知的情况下，作假设检验：

$$H_0 : \sigma_1^2 = \sigma_2^2, H_1 : \sigma_1^2 \neq \sigma_2^2$$

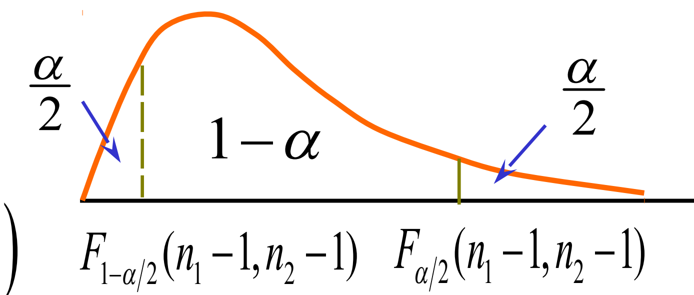
$$\because \frac{\dfrac{(n_1 - 1)S_1^2}{\sigma_1^2} / (n_1 - 1)}{\dfrac{(n_2 - 1)S_2^2}{\sigma_2^2} / (n_2 - 1)} \overset{H_0\text{为真}}{=} \frac{S_1^2}{S_2^2}$$

$$\overset{H_0\text{为真}}{\sim} F\left(n_1 - 1, n_2 - 1\right)$$

$H_0:\sigma_1^2=\sigma_2^2, H_1:\sigma_1^2\neq\sigma_2^2$ 设μ_1,μ_2未知

取检验统计量：

$$F=\frac{S_1^2}{S_2^2}\overset{H_0\text{为真}}{\sim}F(n_1-1,n_2-1)$$



H_0的拒绝域为：

$$\boxed{\frac{S_1^2}{S_2^2}\le F_{1-\alpha/2}(n_1-1,n_2-1),\text{或}\ \frac{S_1^2}{S_2^2}\ge F_{\alpha/2}(n_1-1,n_2-1)}$$

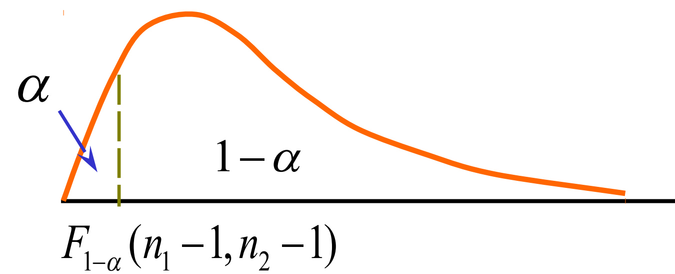
称此检验为F检验法

$$P\_=2*\min\{P[F(n_1-1,n_2-1)\le f_0],P[F(n_1-1,n_2-1)\ge f_0]\}$$

记$p_0=P\{F(n_1-1,n_2-1)\le f_0\}$，则 $P\_=2\min\{p_0,1-p_0\}$

$$H_0:\sigma_1^2 \ge \sigma_2^2,\ H_1:\sigma_1^2 < \sigma_2^2$$ 设μ_1,μ_2未知

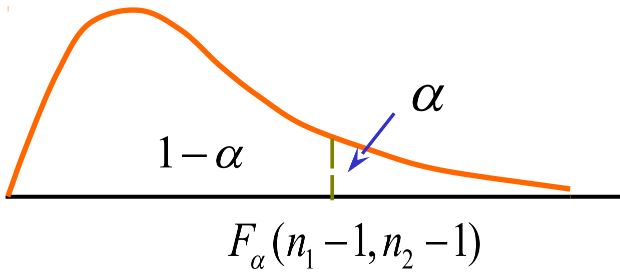
取检验统计量：$F=\dfrac{S_1^2}{S_2^2}$



H_0的拒绝域为：$\boxed{\dfrac{S_1^2}{S_2^2} \le F_{1-\alpha}(n_1-1,n_2-1)}$

$$P\_=P(F(n_1-1,n_2-1)\le f_0)$$

$H_0: \sigma_1^2 \le \sigma_2^2, H_1: \sigma_1^2 > \sigma_2^2$ 设μ_1, μ_2未知



取检验统计量： $F = \dfrac{S_1^2}{S_2^2}$

H_0**的拒绝域为：** $\boxed{\dfrac{S_1^2}{S_2^2} \ge F_\alpha(n_1-1, n_2-1)}$

$$P\_ = P(F(n_1-1, n_2-1) \ge f_0)$$

例 7：两台机床生产同一个型号的滚珠，从甲机床生产的滚珠中抽取 8 个，从乙机床生产的滚珠中抽取 9 个，测得这些滚珠的直径（毫米）如下：

甲机床	15.0	14.8	15.2	15.4	14.9	15.1	15.2	14.8	
乙机床	15.2	15.0	14.8	15.1	14.6	14.8	15.1	14.5	15.0

设两机床生产的滚珠直径分别为X,Y,且

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

(1) 检验假设$H_0:\sigma_1^2=\sigma_2^2, H_1:\sigma_1^2\neq\sigma_2^2(\alpha=0.1)$;

(2) 检验假设$H_0:\mu_1\leq\mu_2, H_1:\mu_1>\mu_2(\alpha=0.1)$;

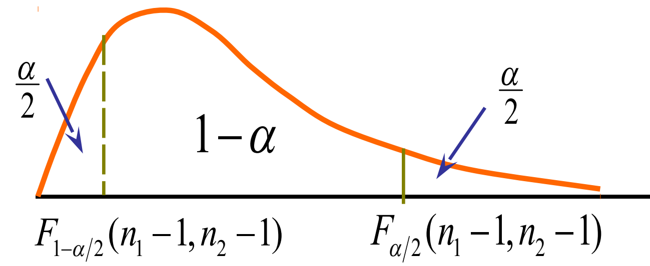
(3) 检验假设$H_0:\mu_1=\mu_2, H_1:\mu_1\neq\mu_2(\alpha=0.1)$。

解：(1) 当μ_1,μ_2未知时，检验 $H_0: \sigma_1^2=\sigma_2^2, H_1: \sigma_1^2 \neq \sigma_2^2$

取检验统计量：$F=\dfrac{S_1^2}{S_2^2} \overset{H_0\text{为真}}{\sim} F(n_1-1, n_2-1)$

H_0的拒绝域为：$\dfrac{S_1^2}{S_2^2} \le F_{1-\alpha/2}(n_1-1, n_2-1)$,

或：$\dfrac{S_1^2}{S_2^2} \ge F_{\alpha/2}(n_1-1, n_2-1)$



查表得：$F_{0.05}(7,8)=3.50, F_{0.95}(7,8)=\dfrac{1}{F_{0.05}(8,7)}=\dfrac{1}{3.73}=0.268$

本题中 $n_1=8$, $\bar{x}=15.05$, $s_1^2=0.0457$；$n_2=9$, $\bar{y}=14.9$, $s_2^2=0.0575$

计算得：$0.268<\dfrac{s_1^2}{s_2^2}=0.795<3.50$

故接受原假设，认为方差没有显著差异。

两总体方差没有显著差异，称为方差齐性。
注意，只有在此前提下比较均值才有意义。

也可用P值检验! $f_0=\dfrac{s_1^2}{s_2^2}=0.795$

H_0**的拒绝域为:** $\dfrac{S_1^2}{S_2^2}\le F_{1-\alpha/2}(n_1-1,n_2-1)$, **或** $\dfrac{S_1^2}{S_2^2}\ge F_{\alpha/2}(n_1-1,n_2-1)$

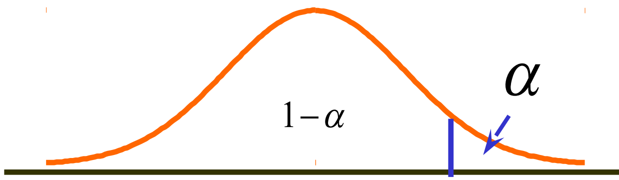
$$
\begin{aligned}
p_0 &= P(F\le f_0)=P(F(8-1,9-1)\le 0.795)\\
&=1-P(F(7,8)>0.795)=1-\textbf{F.DIST.RT(0.795,7,8)}\\
&=1-0.612395=0.387605
\end{aligned}
$$

$$P\_=2\min\{p_0,1-p_0)=2\min\{0.387605,0.612395)=0.77521$$

因为$P_>\alpha=0.1$,故接受原假设，认为方差没有显著差异。

$$n_1=8,\ \ \bar{x}=15.05,\ \ S_1^2=0.0457;\ \ n_2=9,\ \ \bar{y}=14.9,\ \ S_2^2=0.0575$$

$(2)\ H_0:\mu_1\le\mu_2,\ H_1:\mu_1>\mu_2$



H_0的拒绝域为：$\dfrac{\bar{X}-\bar{Y}}{S_w\sqrt{\dfrac{1}{n_1}+\dfrac{1}{n_2}}}\ge t_\alpha(n_1+n_2-2)$

$$t_{0.1}(15)=1.3406,S_w=0.228,\sqrt{\frac{1}{n_1}+\frac{1}{n_2}}=0.486$$

计算得：$\dfrac{\bar{X}-\bar{Y}}{S_w\sqrt{\dfrac{1}{n_1}+\dfrac{1}{n_2}}}=1.354>1.3406$，从而拒绝原假设。

即认为A机床生产的滚珠直径大于B机床生产的滚珠直径。

或用P值判断：

$$P\_=P(t\ge t_0)=P(t(15)\ge1.354)$$

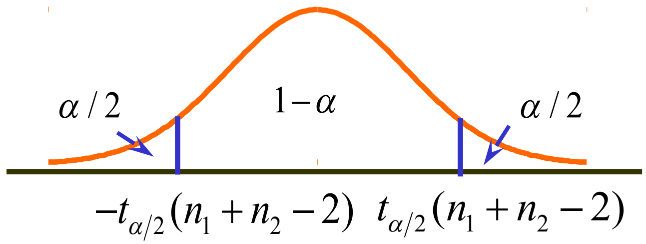
$$=\mathbf{T.DIST.RT(1.354,15)}=0.097893$$

$$<\alpha=0.1$$

从而拒绝原假设。

$$n_1 = 8,\ \bar{x} = 15.05,\ S_1^2 = 0.0457;\ n_2 = 9,\ \bar{y} = 14.9,\ S_2^2 = 0.0575$$

(3) $H_0: \mu_1 = \mu_2, H_1: \mu_1 \neq \mu_2$



H_0的拒绝域为：$\dfrac{|\bar{X}-\bar{Y}|}{S_w\sqrt{\dfrac{1}{n_1}+\dfrac{1}{n_2}}} \geq t_{\frac{\alpha}{2}}(n_1+n_2-2)$

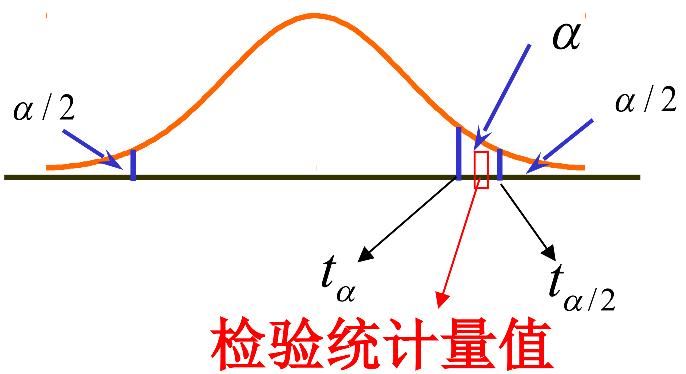
计算得：$\dfrac{|\bar{X}-\bar{Y}|}{S_w\sqrt{\dfrac{1}{n_1}+\dfrac{1}{n_2}}} = 1.354 < t_{0.05}(15) = 1.7531$，从而接受原假设。

即认为A机床生产的滚珠直径与B机床生产的并无显著差异。

或用P值判断：

$$P\_ = P(|t| \geq |t_0|) = 2P(t \geq |t_0|)$$
$$= 2P(t(15) \geq 1.354)$$
$$= 2*\mathbf{TDIST(1.354,15,1)}$$
$$= 2*0.097893 \ > \alpha = 0.1$$

从而接受原假设。



正态总体中参数的双侧置信区间估计与双边假设检验比较

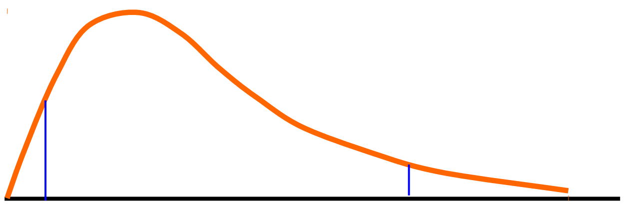
	待估参数	原假设 H_0	枢轴量 G	检验统计量	分布	置信区间	拒绝域
个正态总体	μ （σ^2已知）	$\mu=\mu_0$ （σ^2已知）	$\dfrac{\bar{X}-\mu}{\sigma/\sqrt{n}}$	$\dfrac{\bar{X}-\mu_0}{\sigma/\sqrt{n}}$	$N(0,1)$	$\dfrac{\lvert\bar{X}-\mu\rvert}{\sigma/\sqrt{n}}<z_{\alpha/2}$	$\dfrac{\lvert\bar{X}-\mu_0\rvert}{\sigma/\sqrt{n}}\geq z_{\alpha/2}$
	μ （σ^2未知）	$\mu=\mu_0$ （σ^2未知）	$\dfrac{\bar{X}-\mu}{S/\sqrt{n}}$	$\dfrac{\bar{X}-\mu_0}{S/\sqrt{n}}$	$t(n-1)$	$\dfrac{\lvert\bar{X}-\mu\rvert}{S/\sqrt{n}}<t_{\alpha/2}(n-1)$	$\dfrac{\lvert\bar{X}-\mu_0\rvert}{S/\sqrt{n}}\geq t_{\alpha/2}(n-1)$
	σ^2 (μ未知)	$\sigma^2=\sigma_0^2$ (μ未知)	$\dfrac{(n-1)S^2}{\sigma^2}$	$\dfrac{(n-1)S^2}{\sigma_0^2}$	$\chi^2(n-1)$	$\chi^2_{1-\alpha/2}(n-1)$ $<(n-1)S^2/\sigma^2<$ $\chi^2_{\alpha/2}(n-1)$	$\dfrac{(n-1)S^2}{\sigma_0^2}\leq\chi^2_{1-\alpha/2}(n-1)$ 或$\dfrac{(n-1)S^2}{\sigma_0^2}\geq\chi^2_{\alpha/2}(n-1)$
两个正态总体	$\mu_1-\mu_2$ $(\sigma_1^2=\sigma_2^2=\sigma^2)$	$\mu_1=\mu_2$ $(\sigma_1^2=\sigma_2^2=\sigma^2)$	$\dfrac{(\bar{X}-\bar{Y})-(\mu_1-\mu_2)}{S_w\sqrt{\frac{1}{n_1}+\frac{1}{n_2}}}$	$\dfrac{\bar{X}-\bar{Y}}{S_w\sqrt{\frac{1}{n_1}+\frac{1}{n_2}}}$	$t(n_1+n_2-2)$	$\dfrac{\lvert(\bar{X}-\bar{Y})-(\mu_1-\mu_2)\rvert}{S_w\sqrt{\frac{1}{n_1}+\frac{1}{n_2}}}$ $<t_{\alpha/2}(n_1+n_2-2)$	$\dfrac{\lvert\bar{X}-\bar{Y}\rvert}{S_w\sqrt{\frac{1}{n_1}+\frac{1}{n_2}}}$ $\geq t_{\alpha/2}(n_1+n_2-2)$
	$\dfrac{\sigma_1^2}{\sigma_2^2}$	$\sigma_1^2=\sigma_2^2$	$\dfrac{S_1^2}{S_2^2}\Big/\dfrac{\sigma_1^2}{\sigma_2^2}$	$\dfrac{S_1^2}{S_2^2}$	$F(n_1-1,n_2-1)$	$F_{1-\alpha/2}(n_1-1,n_2-1)$ $<\dfrac{S_1^2}{S_2^2}\Big/\dfrac{\sigma_1^2}{\sigma_2^2}<$ $F_{\alpha/2}(n_1-1,n_2-1)$	$\dfrac{S_1^2}{S_2^2}\leq F_{1-\alpha/2}(n_1-1,n_2-1)$ 或$\dfrac{S_1^2}{S_2^2}\geq F_{\alpha/2}(n_1-1,n_2-1)$

例8: 自两个独立总体$N(\mu_1,\sigma_1^2)$与$N(\mu_2,\sigma_2^2)$中抽取了n_1、n_2个数据,并已计算得:$\bar{x},s_1$;$\bar{y},s_2$,在水平α下,检验

$$H_0:\sigma_1=2\sigma_2 \qquad H_1:\sigma_1\neq 2\sigma_2$$

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

$\therefore$取检验统计量: $F=\dfrac{S_1^2}{4S_2^2}$



$$H_0\text{拒绝域为}F=\frac{S_1^2}{4S_2^2}\leq F_{1-\alpha/2}(n_1-1,n_2-1)$$

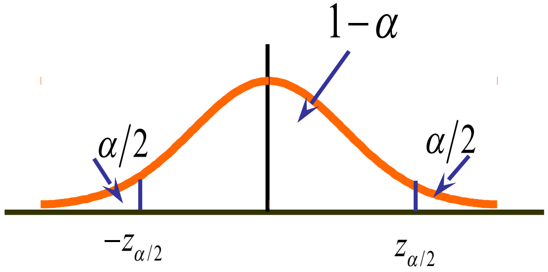
$$\text{或}F=\frac{S_1^2}{4S_2^2}\geq F_{\alpha/2}(n_1-1,n_2-1)$$

§ 8. 4 假设检验与区间估计

若总体$X \sim N(\mu,\sigma^2)$，μ未知，σ^2已知，对于样本

$X_1, X_2, \ldots, X_n$,置信度设为$1-\alpha$，则μ的置信区间为：

$$\bar{X} - \frac{\sigma}{\sqrt{n}} z_{\alpha/2} < \mu < \bar{X} + \frac{\sigma}{\sqrt{n}} z_{\alpha/2}$$



即：$P_{\mu}\left\{\frac{|\bar{X}-\mu|}{\sigma/\sqrt{n}} < z_{\alpha/2}\right\} = 1-\alpha$

对于假设检验问题$H_0: \mu = \mu_0$, $H_1: \mu \neq \mu_0$,显著性水平为α，

H_0 的拒绝域为:$\frac{|\bar{X}-\mu_0|}{\sigma/\sqrt{n}} \geq z_{\alpha/2}$，　即 $P_{\mu_0}\left\{\frac{|\bar{X}-\mu_0|}{\sigma/\sqrt{n}} \geq z_{\alpha/2}\right\} = \alpha$

即拒绝域可以这样得到：将置信区间不等号反向，将原假设成立时的值代入到参数中即可。

双侧置信限与双边假设检验的关系：

一般，若假设检验问题$H_0: \theta = \theta_0$，$H_1: \theta \neq \theta_0$的显著水平为$\alpha$的接受域能等价地写成$\hat{\theta}_L < \theta_0 < \hat{\theta}_U$，那么（$\hat{\theta}_L$，$\hat{\theta}_U$）是参数$\theta$的置信水平为$1-\alpha$的置信区间.

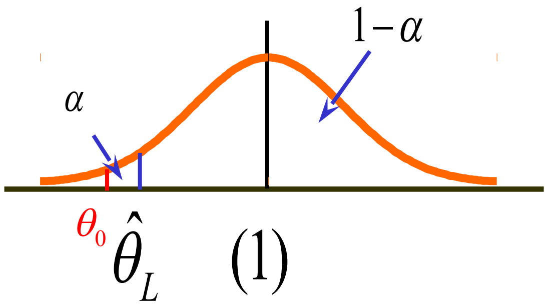
反之，若（$\hat{\theta}_L$，$\hat{\theta}_U$）是θ的置信水平为$1-\alpha$的置信区间，则当$\theta_0 \in$（$\hat{\theta}_L$，$\hat{\theta}_U$）时，在α水平下接受双边检验$H_0: \theta = \theta_0$，$H_1: \theta \neq \theta_0$中的原假设$H_0$,且检验的拒绝域为 $\theta_0 \leq \hat{\theta}_L$或$\theta_0 \geq \hat{\theta}_U$.

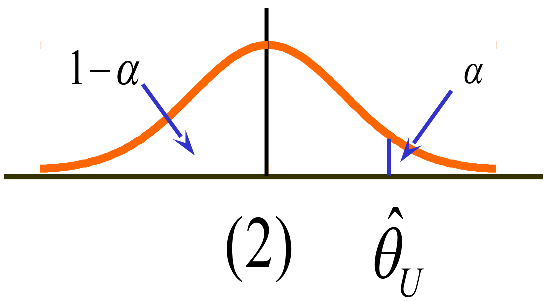
*单侧置信限与单边假设检验的关系：

（1）若$\hat{\theta}_L$是θ的置信水平为$1-\alpha$的单侧置信下限，对右边检验“$H_0:\theta\le\theta_0, H_1:\theta>\theta_0$”中的$\theta_0$,当$\theta_0>\hat{\theta}_L$时，则接受$H_0$；反之，当$\theta_0\le\hat{\theta}_L$时，拒绝原假设.

$$P(\theta>\hat{\theta}_L)\ge 1-\alpha,\text{当}\theta_0\le\hat{\theta}_L\text{时}P(\theta>\theta_0)\ge 1-\alpha, P(\theta\le\theta_0)<\alpha$$

（2）若$\hat{\theta}_U$是θ的置信水平为$1-\alpha$的单侧置信上限，对左边检验“$H_0:\theta\ge\theta_0, H_1:\theta<\theta_0$”中的$\theta_0$,当$\theta_0<\hat{\theta}_U$时，接受$H_0$；反之，当$\theta_0\ge\hat{\theta}_U$时，拒绝原假设.





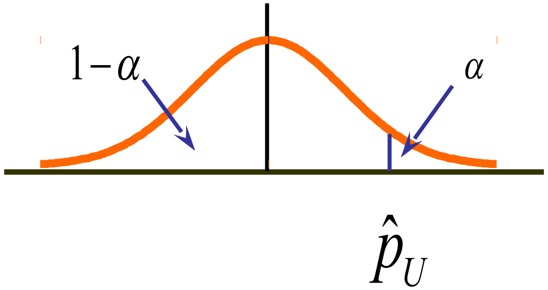
*例9: 为了研究某种止痛药的副作用，调查了服用该种止痛药的440名患者，发现有23名出现了“反症状”，那么是否有足够的理由说明在服用该种止痛药的病人中，出现“反症状”的比例低于10%?

解 设总体$X=\begin{cases}1，出现反症状\\0，不出现反症状\end{cases}$，$X\sim B(1;p)$

由题意知考虑左边检验 $H_0: p\geq 10\%$，$H_1: p<10\%$

根据7.5节有关0-1分布$B(1;p)$中参数p的区间估计的理论，得p的置信水平为$1-\alpha$的单侧置信上限为

$$\hat{p}_U = \frac{1}{2a}\left(-b+\sqrt{b^2-4ac}\right)$$



其中$a = n + z_\alpha^2, b = -(2n\overline{X} + z_\alpha^2), C = n\overline{X}^2.$

将已知资料$n = 440$，$\overline{x} = 23/440 = 0.0523,$

并取$\alpha = 0.05$，代入计算得 $\hat{p}_U \approx 0.0726$

由于$p_0 = 0.1 > \hat{p}_U = 0.0726$,因此，作出拒绝原假设的判断，即认为服用该种止痛药的病人中，出现“ 反症状”的比例低于10%

§8.5 拟合优度检验

前面介绍的各种检验法都是在总体服从正态分布前提下，对参数进行假设检验的。实际中可能遇到这样的情形，总体服从哪种理论分布并不知道，要求我们直接对总体分布提出一个假设。

例如，要检验一个计算机程序产生随机数是否正常。运行该程序产生 100 个 0 ～ 9 之间的数字，观察整数的频数

如下表。那么以 0.05 的显著性水平，是否有充分的理由相信该程序能均匀产生 0 ～ 9 间的整数吗？

整数	0	1	2	3	4	5	6	7	8	9
频数	11	8	7	7	10	10	8	11	14	14

例如，从 1500 到 1931 年的 432 年间，每年爆发战争的次数可以看作一个随机变量，据统计，这 432 年间共爆发了 299 次战争，具体数据如下：

战争次数 X	0	1	2	3	≥ 4
发生 X 次战争的年数	223	142	48	15	4

通常假设每年爆发战争的次数服从泊松分布。那么上面的数据是否有充分的理由推翻每年爆发战争的次数服从泊松分布假设?

（一）皮尔逊拟合优度χ^2检验

它是在总体 X 的分布未知时，根据来自总体的样本，检验关于总体分布的假设的一种检验方法。

先提出假设

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$F(x)$。

注1：若总体X为离散型，则H_0相当于

H_0：总体X的分布律为$P\{X=x_i\}=p_i, i=1,2,\ldots$。

若总体X为连续型，则H_0相当于

H_0：总体X的概率密度为$f(x)$。

注2：当H_0中的总体X的分布函数$F(x)$含有未知参数时，要先用样本求出参数的**极大似然估计**,以估计值为参数值。

χ^2拟合检验法的原理及步骤如下：

(先设H_0中所假设的X的分布函数$F(x)$不含未知参数。)

A_i	n_i	n_i/n	p_i
A_1	n_1	n_1/n	p_1
A_2	n_2	n_2/n	p_2
$\cdots$	$\cdots$	$\cdots$	$\cdots$
A_k	n_k	n_k/n	p_k

1. 从总体取得样本容量为n的样本，将其分成k个两两不相交的子集$A_1,\ldots,A_k$.

2. 以$n_i(i=1,\ldots,k)$记样本观察值$x_1,\ldots,x_n$中落在A_i的个数. 则在n次试验中A_i发生的频率为 n_i/n.
   且有 $n_1+\ldots+n_k = n$

3. 当H_0为真时，容易计算事件A_i发生的概率
   $p_i = P_{H_0}(A_i), i=1,\ldots,k.$ np_i称为**理论频数**.
   而 n_i 是**实际频数**，且有 $n_i \sim B(n,p_i)$

4. H_0的拒绝域形式为：$\displaystyle\sum_{i=1}^{k}\frac{n}{p_i}\left(\frac{n_i}{n}-p_i\right)^2 \geq c.$

$\because n_i \sim B(n, p_i)$

$\therefore$ 当n足够大时 $n_i \overset{近似}{\sim} N(np_i, np_i q_i)$

$$\Rightarrow \frac{n_i - np_i}{\sqrt{np_i q_i}} \overset{近似}{\sim} N(0,1) \quad \Rightarrow \frac{(n_i - np_i)^2}{np_i q_i} \overset{近似}{\sim} \chi^2(1)$$

$$\because \sum_{i=1}^{k}(n_i - np_i) = n - n = 0 \quad \therefore \sum_{i=1}^{k}\frac{(n_i - np_i)^2}{np_i q_i} \overset{近似}{\sim} \chi^2(k-1)$$

若p_i较小时，$q_i = 1 - p_i \approx 1$

$$\therefore \sum_{i=1}^{k}\frac{(n_i - np_i)^2}{np_i} = \sum_{i=1}^{k}\frac{n}{p_i}\left(\frac{n_i}{n} - p_i\right)^2 \overset{近似}{\sim} \chi^2(k-1)$$

化简： $\displaystyle\sum_{i=1}^{k}\frac{(n_i-np_i)^2}{np_i}$

$$=\sum_{i=1}^{k}\frac{n_i^2-2n_inp_i+(np_i)^2}{np_i}$$

$$=\sum_{i=1}^{k}\frac{n_i^2}{np_i}-2\sum_{i=1}^{k}n_i+n\sum_{i=1}^{k}p_i$$

$$=\sum_{i=1}^{k}\frac{n_i^2}{np_i}-2\cdot n+n\cdot 1 \quad =\sum_{i=1}^{k}\frac{n_i^2}{np_i}-n$$

定理. 若n充分大($n \geq 50$)，当H_0为真时，则统计量

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

近似服从$\chi^2(k-1)$分布。因此检验拒绝域为

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

注1：当H_0中所假设的X的分布函数$F(x)$含有r个未知参数时，

先用样本求出未知参数的极大似然估计，以估计值为参数值，

求出p_i的估计值$\hat{p}_i = \hat{P}_{H_0}(A_i)$，于是检验统计量为

$$\chi^2 = \sum_{i=1}^{k} \frac{n_i^2}{n\hat{p}_i} - n \overset{\text{近似}}{\sim} \chi^2(k-r-1),$$

因此检验拒绝域为 $\chi^2 \geq \chi_\alpha^2(k-r-1)$

在显著性水平α下拒绝域为

$$\chi^2=\sum_{i=1}^{k}\frac{n_i^2}{np_i}-n\geq\chi_\alpha^2(k-1),$$ （没有参数需要估计）

$$\chi^2=\sum_{i=1}^{k}\frac{n_i^2}{n\hat{p}_i}-n\geq\chi_\alpha^2(k-r-1),$$ （有r个参数需要估计）

注2: χ^2拟合检验使用时必须注意n要足够大，np_i(或$n\hat{p}_i$)不能太小。根据实践，要求$n\geq 50$，np_i (或 $n\hat{p}_i$) ≥ 5，否则应适当合并A_i，以满足要求。

例 1：从 1500 到 1931 年的 432 年间，每年爆发战争的次数可以看作一个随机变量，据统计，这 432 年间共爆发了 299 次战争，具体数据如下：

年内战争次数 X	0	1	2	3	≥ 4
发生 X 次战争的年数	223	142	48	15	4

通常假设每年爆发战争的次数服从泊松分布。那么上面的数据是否有充分的理由推翻每年爆发战争的次数服从泊松分布的假设？（取 α=0.05 ）

解：$H_0 : X \sim P(\lambda)$，λ未知，$\hat{\lambda}_L = \bar{x} = 299/432 = 0.69.$

$$\hat{p}_i = P(X=i) = \frac{\hat{\lambda}^i e^{-\hat{\lambda}}}{i!}, i = 0,1,2,3, \quad \hat{p}_4 = \sum_{j=4}^{\infty} \frac{\hat{\lambda}^j e^{-\hat{\lambda}}}{j!}, \quad n = 432$$

战争次数 X	0	1	2	3	>=4
实测频数 n_i	223	142	48	15	4
概率估计 $\hat{p}_i$	0.502	0.346	0.119	0.027	0.006
理论频数 $n\hat{p}_i$	217	149	51	12	3

战争次数	X	0	1	2	3	>=4
实测频数	n_i	223	142	48	15	4
概率估计	$\hat{p}_i$	0.502	0.346	0.119	0.027	0.006
理论频数	$n\hat{p}_i$	217	149	51		

12 3

15

检验统计量的观察值为

$$\chi^2 = \sum_{i=1}^{k} \frac{n_i^{\ 2}}{n\hat{p}_i} - n = \frac{223^2}{217} + \frac{142^2}{149} + \frac{48^2}{51} + \frac{19^2}{15} - 432 = 1.74$$

H_0的拒绝域为：$\chi^2 \geq \chi^2_{\alpha}(k-r-1)$

其中：$n = 432, \alpha = 0.05, k = 4, r = 1$

$$\chi_{\alpha}{}^2(k-r-1) = \chi^2_{0.05}(4-1-1) = 5.991$$

于是， $1.74 < 5.991$,不能拒绝原假设。即认为每年爆发战争的次数服从泊松分布。

例 2　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、这些数据提供充分证据拒绝该理论吗？（取 α= 0.05）

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

$n = 315 + 101 + 108 + 32 = 556$

$k = 4, r = 0, \alpha = 0.05$

$$H_0 : P(X=1) = \frac{9}{16}, P(X=2) = \frac{3}{16}, P(X=3) = \frac{3}{16}, P(X=4) = \frac{1}{16}$$

豆子状态 X	1	2	3	4
实测频数 n_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

$$\chi^2 = \sum_{i=1}^{k} \frac{n_i^2}{np_i} - n = 0.47 < \chi^2_{0.05}(3) = 7.815,$$ 接受 H_0. 即理论正确。

例 3　下面列出了 84 个伊特拉斯坎 (*Etruscan*) 人男子的头颅的最大宽度 (mm)，试检验这些数据是否来自正态总体。（取 α=0. 1 ）

141	148	132	138	154	142	150	146	155	158	150	140
147	148	144	150	149	145	149	158	143	141	144	144
126	140	144	142	141	140	145	135	147	146	141	136
140	146	142	137	148	154	137	139	143	140	131	143
141	149	148	135	148	152	143	144	141	143	147	146
150	132	142	142	143	153	149	146	149	138	142	149
142	137	134	144	146	147	140	142	140	137	152	145

说明：为什么题目要我们检验这些数据来自正态总体呢？

一般，用直方图法粗略了解数据的分布情况。 步骤如下：

1. 找出数据的最小值、最大值为 126 、 158 ，取区间 [124.5, 159.5]，它能覆盖 [126, 158] ;

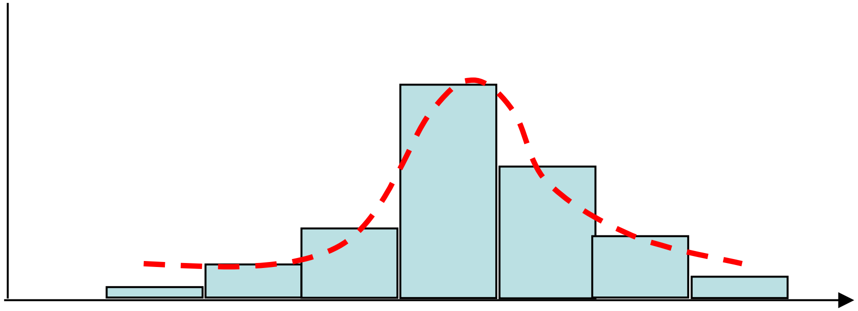
2. 将区间 [124.5, 159.5] 等分为 7 个小区间 , 小区间的长度 $\Delta=(159.5-124.5)/7=5$, Δ 称为组距，小区间的端点称为组限，建立下表：

组　限	频数 n_i	频率　n_i/n	累计频率
124.5-129.5	1	0.0119	0.0119
129.5-134.5	4	0.0476	0.0595
134.5-139.5	10	0.1191	0.1786
139.5-144.5	33	0.3929	0.5715
144.5-149.5	24	0.2857	0.8572
149.5-154.5	9	0.1071	0.9524
154.5-159.5	3	0.0357	1

3. 自左向右在各小区间上作以 $\dfrac{n_i}{n}\Big/\Delta$ 为高的小矩形，如下图，即为直方图。

注：直方图的小区间可以不等长，但小区间的长度不能太大，否则平均化作用突出，淹没了密度的细节部分；也不能太小，否则受随机化影响太大，产生极不规则的形状。

这样，各矩形面积为频率，当 n 充分大时接近于概率。



小矩形面积与密度函数曲线之下该小区间的曲边梯形面积基本相同。根据频率画出密度函数。

从本例的直方图看，有一个峰，中间高，两头低，较对称，和正态总体很相似。于是检验

$H_0: X$的概率密度为

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

H_0中μ,σ^2未知，先求出其**最大似然估计**分别为

$$\hat{\mu}=\overline{X}=143.8,\ \hat{\sigma}_L^2=B_2=6.0^2,$$

此时X的概率密度的估计为

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

A_i	n_i	$\hat{p}_i$	$n\hat{p}_i$	$n_i^2 / n\hat{p}_i$
A_1 $-\infty < x \le 129.5$	1	0.0087	0.73 } 5.09	4.91
A_2 $129.5 < x \le 134.5$	4	0.0519	4.36	
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	24.40
A_6 $149.5 < x \le 154.5$	9	0.1336	11.22 } 14.37	10.02
A_7 $154.5 < x < \infty$	3	0.0375	3.15	$\Sigma=87.67$

$$\chi^2 = \sum_{i=1}^{k} \frac{n_i^2}{n\hat{p}_i} - n = 87.67 - 84 = 3.67$$

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

故在水平 0.1 下接受 H_0，认为数据来自正态总体。

例 4 一台摇奖机是一个圆球形容器，内有 10 个质地均匀的小球，分别标有 0，1，2，…，9 的数码。转动容器让小球随机分布，然后从中掉出一球，其号码为 X。如果摇奖机合格，则 X 的分布律应为

$$P(X=k)=\frac{1}{10}\qquad k=0,1,2,\cdots,9$$

现用这台摇奖机做了 800 次试验，得到如下数据：

号码	0	1	2	3	4	5	6	7	8	9
出现的频数	74	92	83	79	80	73	77	75	76	91

问在 0.05 显著性水平下，能否认为该摇奖机合格？

解 由题意要检验假设

$$H_0: P(X=k)=\frac{1}{10} \qquad k=0,1,2,\cdots,9$$

将已知数据按号码分为 10 组，分组为

$$A_i=\{i\} \qquad i=0,1,2,\cdots,9$$

记 $p_i=P(X\in A_i) \qquad i=0,1,2,\cdots,9$

当 H_0 为真时 $p_i=\dfrac{1}{10}$

计算列于下表中

χ^2 检验计算表

A_i	n_i	p_i	np_i	$\frac{(n_i - np_i)^2}{np_i}$
A_0	74	0.1	80	0.45
A_1	92	0.1	80	1.8
A_2	83	0.1	80	0.1125
A_3	79	0.1	80	0.0125
A_4	80	0.1	80	0
A_5	73	0.1	80	0.6125
A_6	77	0.1	80	0.1125
A_7	75	0.1	80	0.3125
A_8	76	0.1	80	0.2
A_9	71	0.1	80	0.5125
$\sum$				4.125

H_0 的检验统计量及拒绝域：

$$\chi^2 = \sum_{i=0}^{9} \frac{(n_i - np_i)^2}{np_i} \geq \chi_\alpha^2(k - r - 1)$$

其中

$$k = 10 \quad r = 0 \quad \alpha = 0.05 \quad \chi_{0.05}^2(9) = 16.919$$

由于

$$\chi^2 = 4.125 < 16.919 = \chi_{0.05}^2(9)$$

故接受 H_0，即认为摇奖机是合格的。

五、（12 分）在某地区随机抽选了 200 位成年人测量身高 X（单位：厘米），数据如下表所示，已知样本均值 $\bar{x}=170$，样本方差 $s^2=\dfrac{5000}{199}$，检验数据是否来自正态总体。（$\alpha=0.05$）

身高范围	$x\le 162.5$	$162.5<x \le 167.5$	$167.5<x \le 172.5$	$172.5<x \le 177.5$	$x>177.5$
频数 f_i	12	47	70	51	20

三.（12 分）为研究某总体 X 的分布，抽取了容量为 162 的简单随机样本，结果如下表：

取值	1	2	3	≥4
样本个数	111	27	16	8

检验原接受 $H_0: P(X=k)=\dfrac{2}{3^k}, k=1,2,\ldots$。（显著性水平为 $\alpha=0.05$）

Pearsonχ^2检验的缺点：

对于连续性随机变量，检验统计量的取值依赖于区间的划分，影响检验的功效。

适用于离散型随机变量的分布检验！

＊（二）柯尔莫哥洛夫

＊（三）正态 W 检验（ D 检验）

*偏度、峰度检验

偏度、峰度检验法是用于检验正态总体的一种方法。

定义：随机变量X的偏度：

$$v_1 = E\left[\left(\frac{X-E(X)}{\sqrt{D(X)}}\right)^3\right] = \frac{E\left[\left(X-E(X)\right)^3\right]}{(D(X))^{3/2}},$$

峰度：$$v_2 = E\left[\left(\frac{X-E(X)}{\sqrt{D(X)}}\right)^4\right] = \frac{E\left[\left(X-E(X)\right)^4\right]}{(D(X))^2},$$

当$X \sim N(\mu, \sigma^2)$时， $v_1 = 0$， $v_2 = 3.$

于是，将检验 $H_0: X \sim N(\mu, \sigma^2)$ 的问题转化为

检验 $H_0': v_1 = 0, v_2 = 3.$

* 秩和检验

秩和检验是替换两个独立总体均值差的 t 检验的一种非参数方法。

两个独立总体均值差的 t 检验是基于两个总体都服从正态分布，且方差未知但相等的情形。

而实际情况中，往往两个独立总体的分布是未知的，但属于同一类，例如都是连续型总体，密度函数只差一个平移。即

设两总体的概率密度分别为$f_1(x)$，$f_2(x)$，则$f_1(x)=f_2(x-a)$

检验假设 $H_0: a=0, H_1: a\neq 0.$ $\quad H_0:\mu_1=\mu_2, H_1:\mu_1\neq\mu_2.$

$H_0: a=0, H_1: a>0.$ $\Leftrightarrow H_0:\mu_1=\mu_2, H_1:\mu_1>\mu_2.$

$H_0: a=0, H_1: a<0.$ $\quad H_0:\mu_1=\mu_2, H_1:\mu_1<\mu_2.$

若两总体的均值存在，分别记为μ_1，μ_2，则

$$\mu_1=\int_{-\infty}^{\infty} x f_1(x)dx=\int_{-\infty}^{\infty} x f_2(x-a)dx \overset{t=x-a}{=} \int_{-\infty}^{\infty}(t+a)f_2(t)dt=\mu_2+a.$$

复习思考题　8

1. 假设检验的基本思想是什么？其中使用了一条什么原理？
2. 检验的显著性水平 α 的意义是什么？
3. 比较双边、左边和右边检验的拒绝域。
4. 使用 Z 检验法可以进行哪些假设检验？
5. 使用 t 检验法可以进行哪些假设检验？
6. 使用 χ^2 检验法可以进行哪些假设检验？
7. 使用 F 检验法可以进行哪些假设检验？
8. 正态总体期望与方差的区间估计和假设检验两者之间有什么相似之处？
9. 成对数据差的 t 检验适用于哪些特殊场合？
10. 分布拟合的 χ^2 检验的基本步骤是什么？

课件待续！

20/2/15