- 统计方法
  - 描述统计
  - 推论统计
    - 参数估计
    - 假设检验

推论统计内容

- 参数估计
  - 点估计
  - 区间估计
- 假设检验
  - 参数检验
    - Z 检验
    - t 检验
    - F 检验
    - q 检验…
  - 非参数检验
    - 计数资料—— χ^2 检验
    - 计量资料
      - 相关样本
        - 符号检验
        - 符号秩次检验
      - 独立样本
        - 秩和检验
        - 中位数检验

参数估计和假设检验

- 参数估计和假设检验是统计推断的两个组成部分，都是利用样本对总体进行某种推断，但推断的角度不同。参数估计讨论的是用样本统计量估计总体参数的方法。假设检验讨论的是用样本信息去检验对总体参数的某种假设是否成立的程序和方法。

我们关心的是从样本统计值得出的差异能否作出一般性结论（样本是否具有推广性）

——总体参数之间确实存在差异。

在统计学中，进行这种推论的过程称作**假设检验**。

讨论统计值之间差异的问题时，有两种情况：

样本统计量与相应总体参数的差异

两个样本统计量之间的差异

显 著

该样本已基本不属于已知总体

各自代表的两个总体的参数之间确实存在差异

Hypothesis testing

- 包括参数和非参数检验:
  - 参数假设检验（parametric test）：总体的分布形式已知，需要对总体的未知参数进行假设检验。
  - 非参数假设检验（non-parametric test）：对总体分布形式所知甚少，需对未知分布函数的形式及其他特征进行假设检验。

参数假设检验举例

例1: 根据1989年的统计资料，某地女性新生儿的平均体重为3190克。为判断该地1990年的女性新生儿体重与1989年相比有无显著差异，从该地1990年的女性新生儿中随机抽取30人，测得其平均体重为3210克。从样本数据看，1990年女新生儿体重比1989年略高，但这种差异可能是由于抽样的随机性带来的，也许这两年新生儿的体重并没有显著差异。究竟是否存在显著差异？可以先假设这两年新生儿的体重没有显著差异，然后利用样本信息检验这个假设能否成立。这是一个关于总体均值的假设检验问题。

参数假设检验举例

例2：某种大量生产的袋装食品，按规定每袋重量不得少于250克，现从一批该种食品中任意抽取50袋，发现有6袋重量低于250克。若规定食品不符合标准的比例达到5％就不得出厂，问该批食品能否出厂。

可以先假设该批食品的不合格率不超过5％，然后用样本不合格率来检验假设是否正确。这是一个关于总体比例的假设检验问题。

本章主要内容

- 假设检验的原理
- 平均数的显著性检验
- 平均数差异的显著性检验
- 方差的差异检验
- 相关系数的显著性检验
- 比率的显著性检验

第一节 假设检验的原理

- **假设检验的原理：**小概率原理。假设检验的基本思想是概率性质的**反证法**。(不同于纯数学中的反证法)
- **什么是小概率?**
- 概率是0～1之间的一个数，因此小概率就是接近0的一个数
- 著名的英国统计家Ronald Fisher 把20分之1作为标准，也就是0.05，从此0.05或比0.05小的概率都被认为是小概率
- Fisher没有任何深奥的理由解释他为什么选择0.05，只是说他忽然想起来的

什么是小概率原理?

- 小概率原理——发生概率很小的随机事件（小概率事件）在一次实验中几乎是不可能发生的。
- 根据这一原理，可以先假设总体参数的某项取值为真，也就是假设其发生的可能性很大，然后抽取一个样本进行观察，如果样本信息显示出现了与事先假设相反的结果且与原假设差别很大，则说明原来假定的小概率事件在一次实验中发生了，这是一个违背小概率原理的不合理现象，因此有理由怀疑和拒绝原假设；否则不能拒绝原假设。
- 检验中使用的小概率是检验前人为指定的。

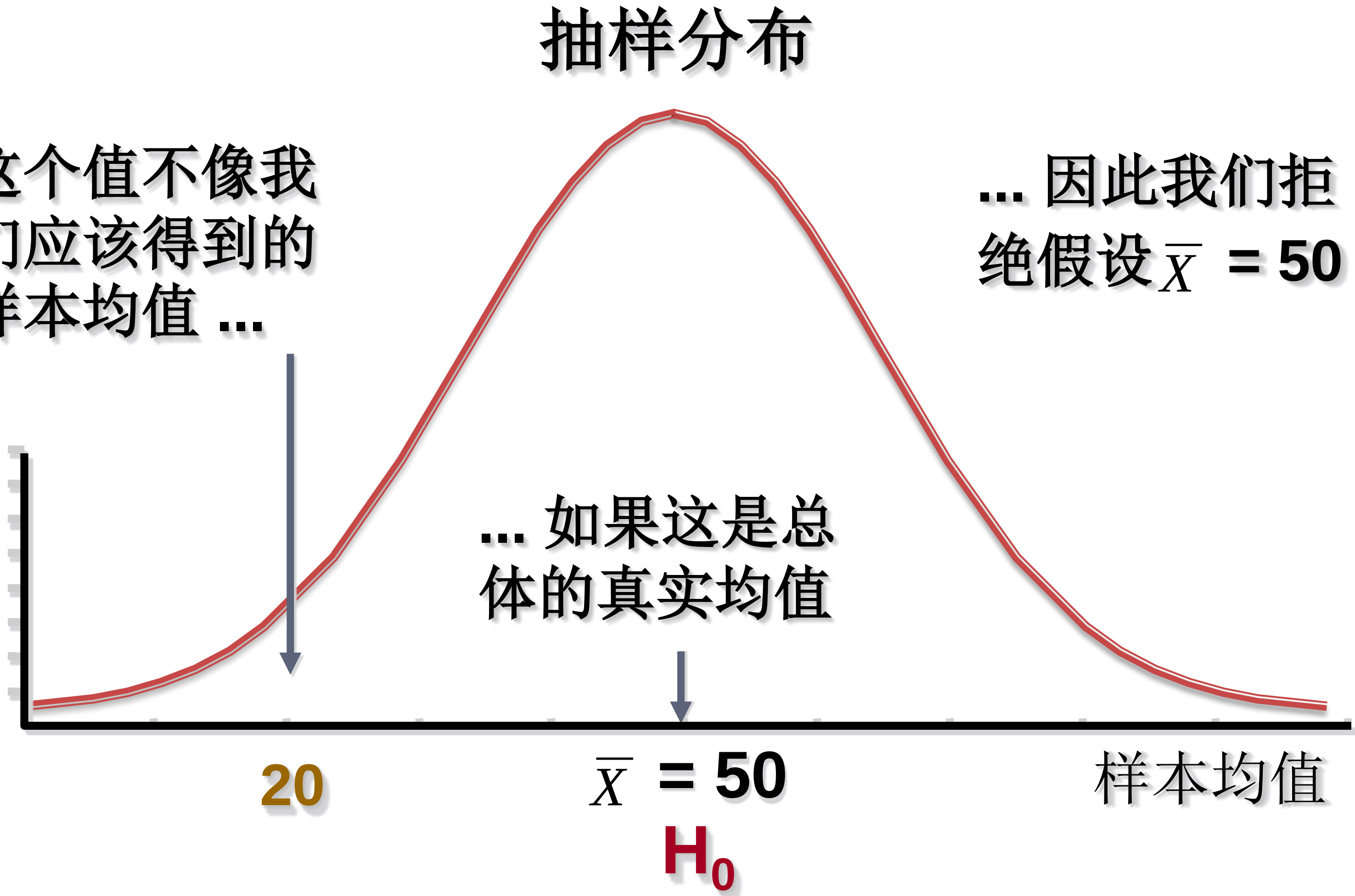
小概率原理举例：

- 某工厂质检部门规定该厂产品次品率不超过4%方能出厂。今从1000件产品中抽出10件，经检验有4件次品，问这批产品是否能出厂?
- 如果假设这批产品的次品率P≤4%，则可计算事件"抽10件产品有4件次品"的出现概率为：

$$P_{10}(4) = C_{10}^{4}(0.04)^{4}(1-0.04)^{6} = 0.00042$$

可见，概率是相当小的，1万次实验中可能出现4次，然而概率如此小的事件，在一次实验中居然发生了，这是不合理的，而不合理的根源在于假设次品率P≤4%，因而认为假设次品率P≤4%是不能成立的，故按质检部门的规定，这批产品不能出厂。

假设检验的基本思想



抽样分布
这个值不像我们应该得到的样本均值 ...
... 因此我们拒绝假设 $\overline{X}$ = 50
... 如果这是总体的真实均值
20
$\overline{X}$ = 50
H_0
样本均值

假设检验的两个特点：

第一，假设检验采用逻辑上的反证法，即为了检验一个假设是否成立，首先假设它是真的，然后对样本进行观察，如果发现出现了不合理现象，则可以认为假设是不合理的，拒绝假设。否则可以认为假设是合理的，接受假设。

第二，假设检验采用的反证法带有概率性质。所谓假设的不合理不是绝对的，而是基于实践中广泛采用的小概率事件几乎不可能发生的原则。至于事件的概率小到什么程度才算是小概率事件，并没有统一的界定标准，而是必须根据具体问题而定。如果一旦判断失误，错误地拒绝原假设会造成巨大损失，那么拒绝原假设的概率就应定的小一些；如果一旦判断失误，错误地接受原假设会造成巨大损失，那么拒绝原假设的概率就应定的大一些。

- 小概率通常用α表示，又称为检验的显著性水平。通常取α＝0.05或α＝0.01，即把概率不超过0.05或0.01的事件当作小概率事件。

原假设和备择假设

- 假设检验中，我们称作为检验对象的待检验假设为原假设或零假设，用H_0表示。原假设的对立假设称为备择假设或备选假设，用H_1表示。
- 例如，设 $\bar{X}_0$为总体均值 $\bar{X}$的某一确定值。

（1）对于总体均值是否等于某一确定值的原假设可以表示为：

H_0： $X= \bar{X}_0$ （如H_0： $\bar{X} = 3190$克）

其对应的备择假设则表示为：

H_1： $X \neq \bar{X}_0$ （如H_1： $\bar{X} \neq 3190$克）

（2）对于总体均值 $\bar{X}$ 是否大于某一确定值 $\bar{X}_0$ 的原假设可以表示为：

H_0: $\bar{X} \geq \bar{X}_0$（如 H_0: $\bar{X} \geq 2000$ 克）

其对应的备择假设则表示为：

H_1: $\bar{X} < \bar{X}_0$（如 H_1: $\bar{X} < 2000$ 克）

（3）对于总体均值 $\bar{X}$ 是否小于某一确定值 $\bar{X}_0$ 的原假设可以表示为：

H_0: $\bar{X} \leq \bar{X}_0$（如 H_0: $\bar{X} \leq 5\%$）

其对应的备择假设则表示为：

H_1: $\bar{X} > \bar{X}_0$（如 H_1: $\bar{X} > 5\%$）

注意：原假设总是有等号：$=$ 或 $\leq$ 或 $\geq$。

双侧检验和单侧检验

- 根据假设的形式不同，假设检验可以分为双侧假设检验和单侧假设检验。
- 若原假设是总体参数等于某一数值，如H_0：$\bar{X}=\bar{X}_0$，即备择假设H_1：$\bar{X}\neq\bar{X}_0$，那么只要$\bar{X}<\bar{X}_0$和$\bar{X}>\bar{X}_0$二者中有一个成立，就可以否定原假设。这种假设检验称为双侧检验。
- 若原假设是总体参数大于等于或小于等于某一数值，如H_0：$\bar{X}\geq\bar{X}_0$（即H_1：$\bar{X}<\bar{X}_0$）；或H_0：$\bar{X}\leq\bar{X}_0$（即H_1：$\bar{X}>\bar{X}_0$），那么对于前者当$\bar{X}<\bar{X}_0$时，对于后者当$\bar{X}>\bar{X}_0$时，可以否定原假设。这种假设检验称为单侧检验。可以分为左侧检验和右侧检验。

双侧检验与单侧检验(假设的形式)

假设	研究的问题（总体均值检验）		
	双侧检验	左侧检验	右侧检验
H_0	$\bar{X} = \bar{X}_0$	$\bar{X} \geq \bar{X}_0$	$\bar{X} \leq \bar{X}_0$
H_1	$\bar{X} \neq \bar{X}_0$	$\bar{X} < \bar{X}_0$	$\bar{X} > \bar{X}_0$

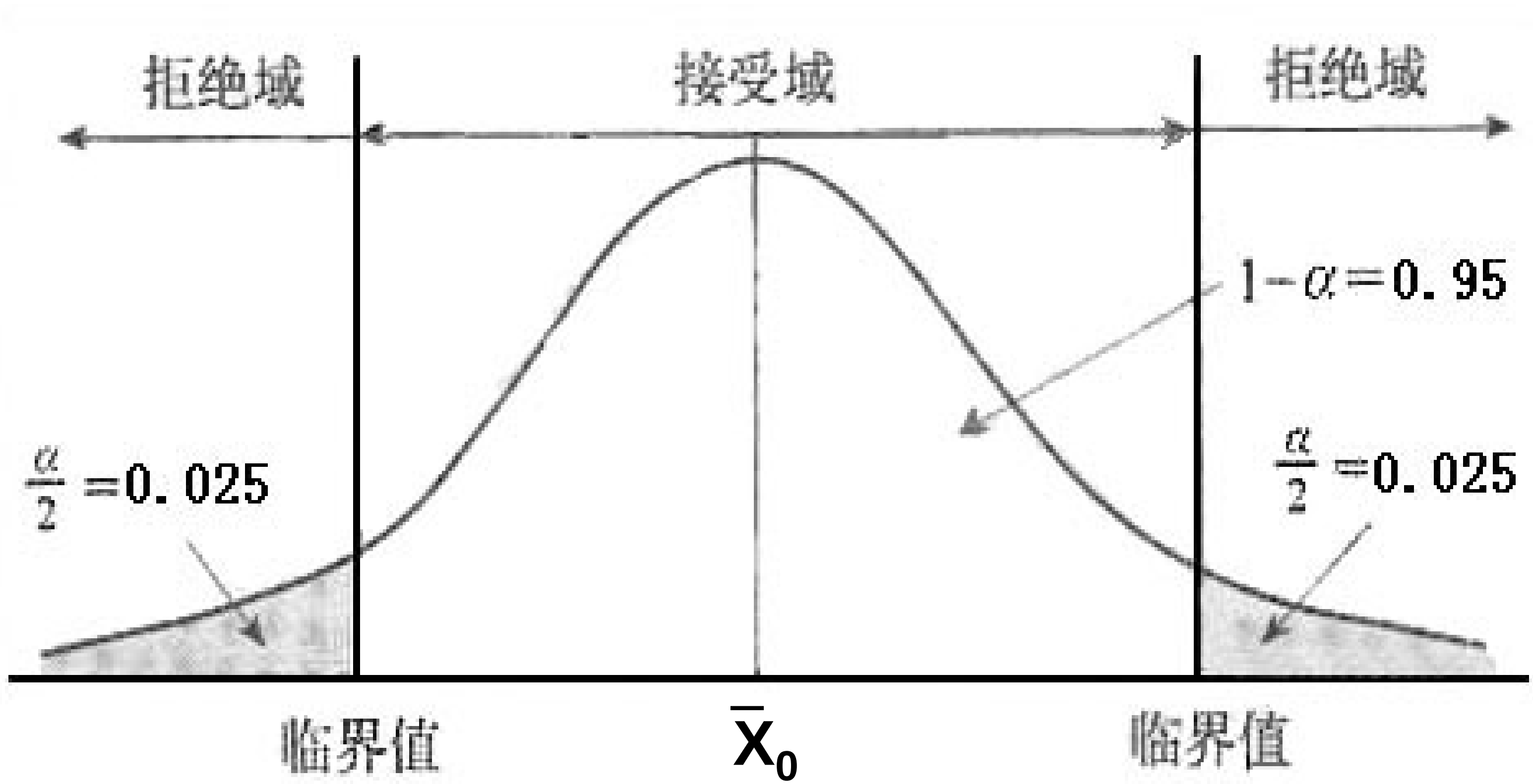
假设检验中的拒绝域和接受域

- 在规定了检验的显著性水平α后，根据容量为n的样本，按照统计量的理论概率分布规律，可以确定据以判断拒绝和接受原假设的检验统计量的临界值。
- 临界值将统计量的所有可能取值区间分为两个互不相交的部分，即原假设的拒绝域和接受域。
- 对于正态总体，总体均值的假设检验可有如下图示：

- 正态总体，总体均值假设检验图示：

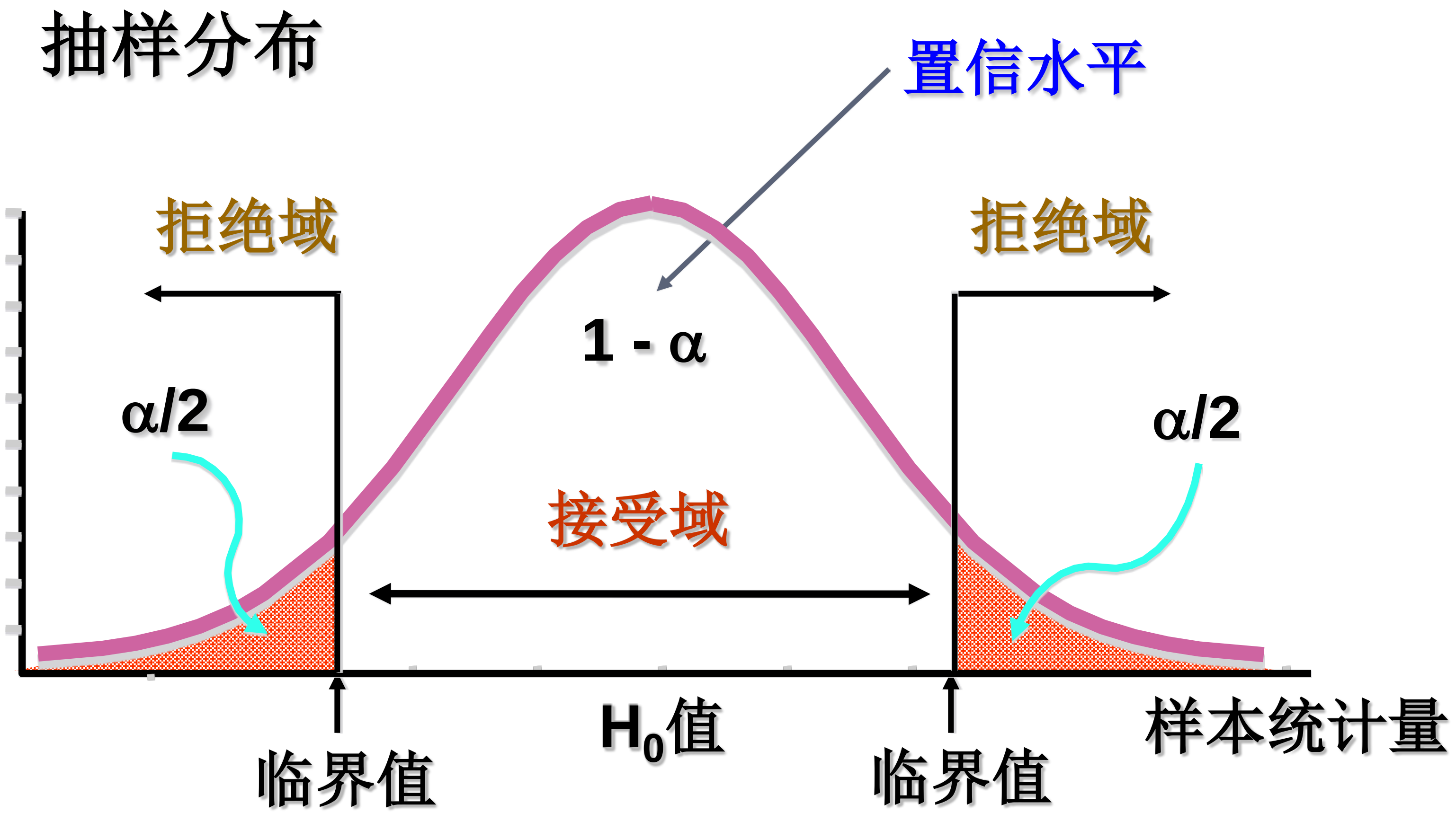
（1） 双侧检验

设H_0： $\bar{X} = \bar{X}_0$， H_1： $\bar{X} \neq \bar{X}_0$，有两个临界值，两个拒绝域，每个拒绝域的面积为α/2。也称双尾检验。

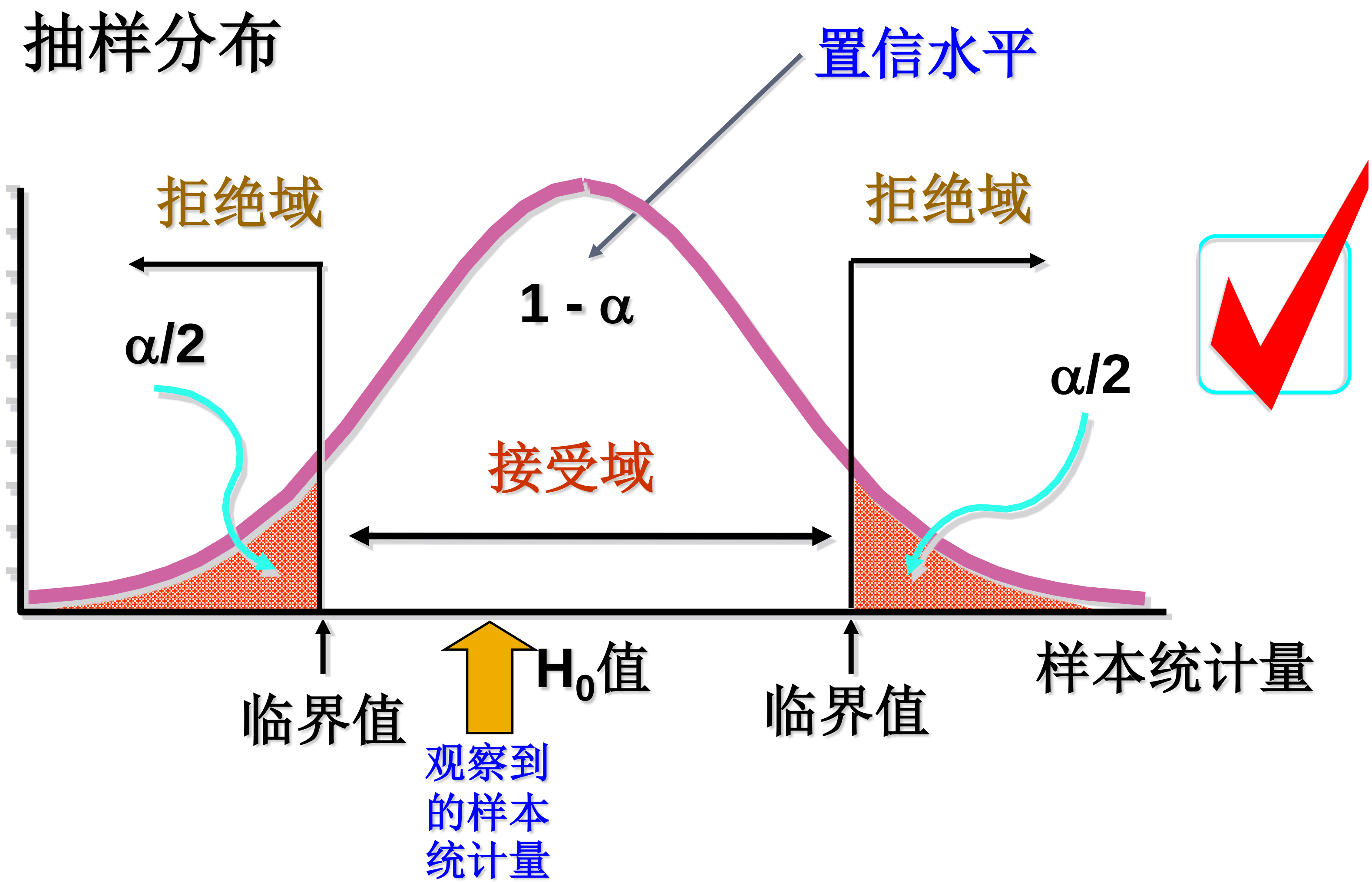


双侧检验示意图

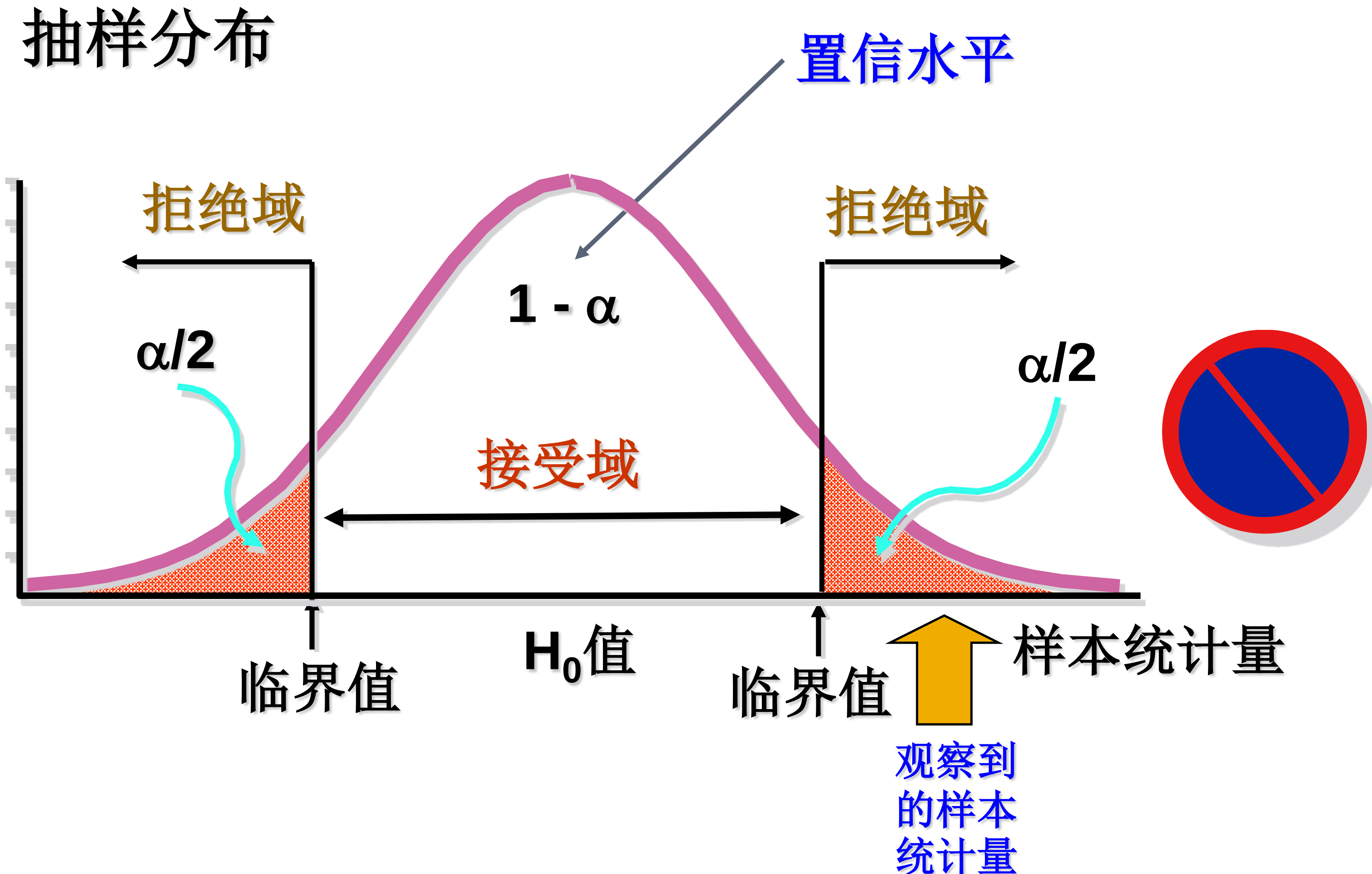
双侧检验示意图（显著性水平与拒绝域）



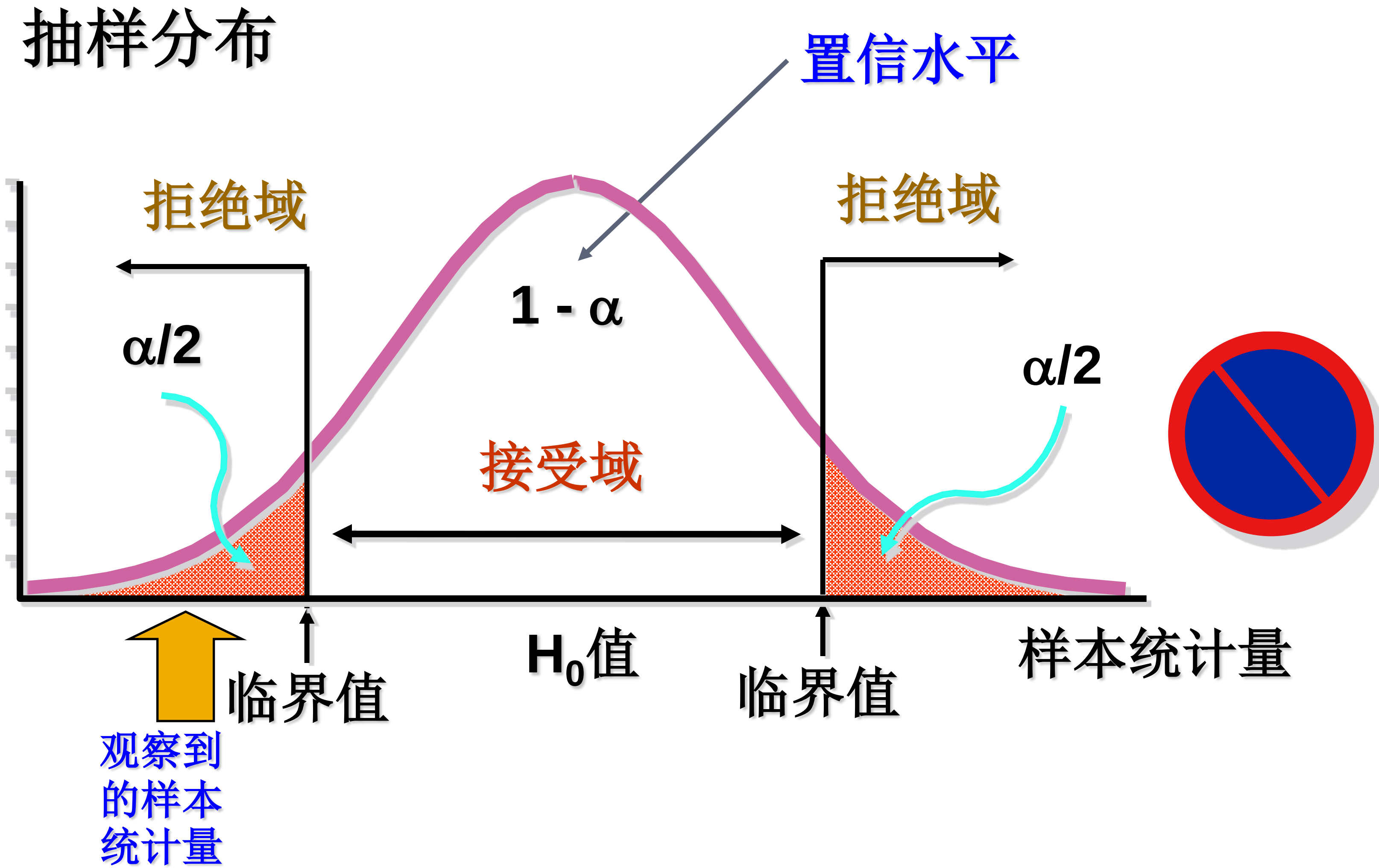
双侧检验示意图（显著性水平与拒绝域）



双侧检验示意图（显著性水平与拒绝域）



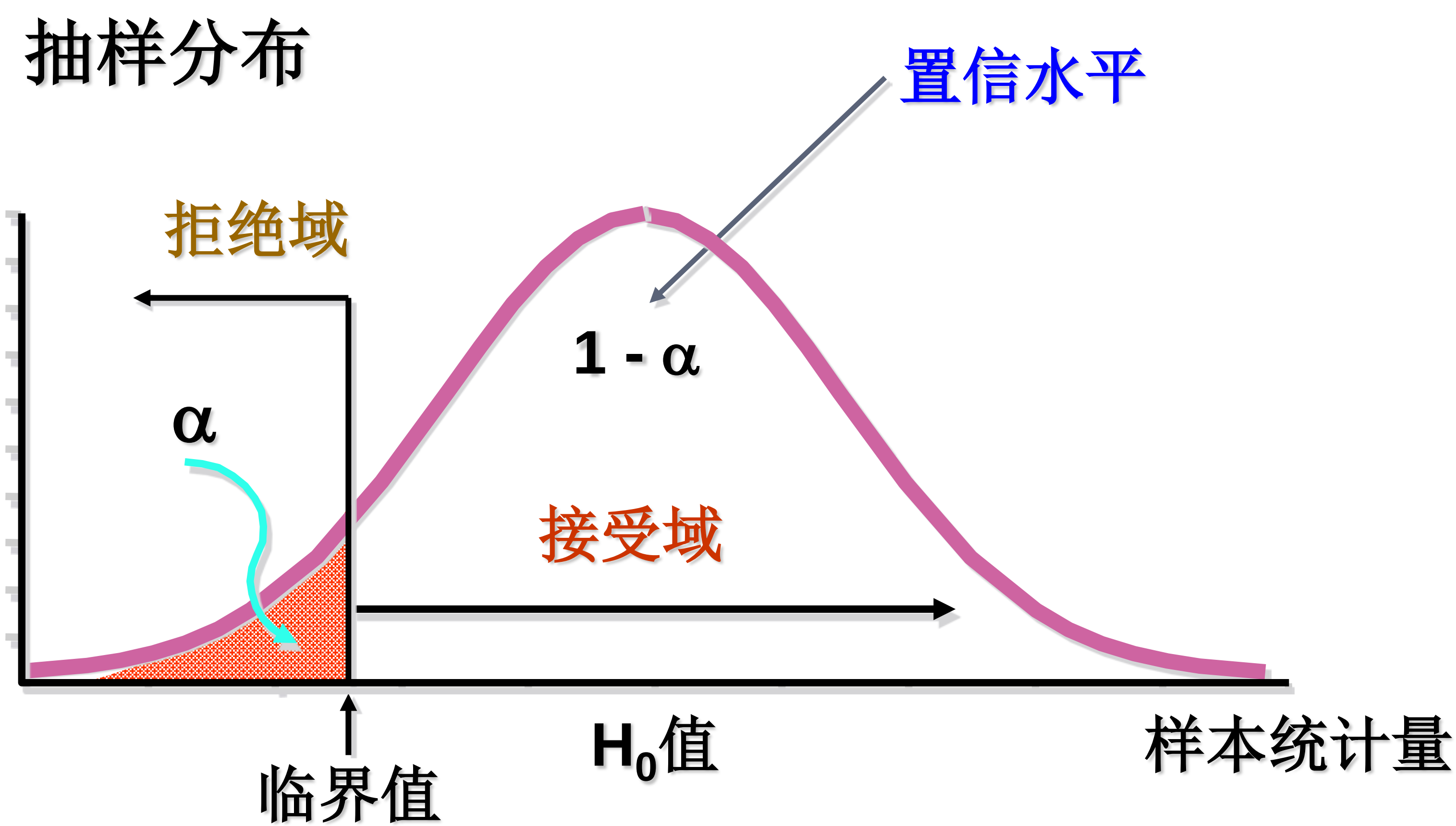
双侧检验示意图
（显著性水平与拒绝域）



（2）单侧检验

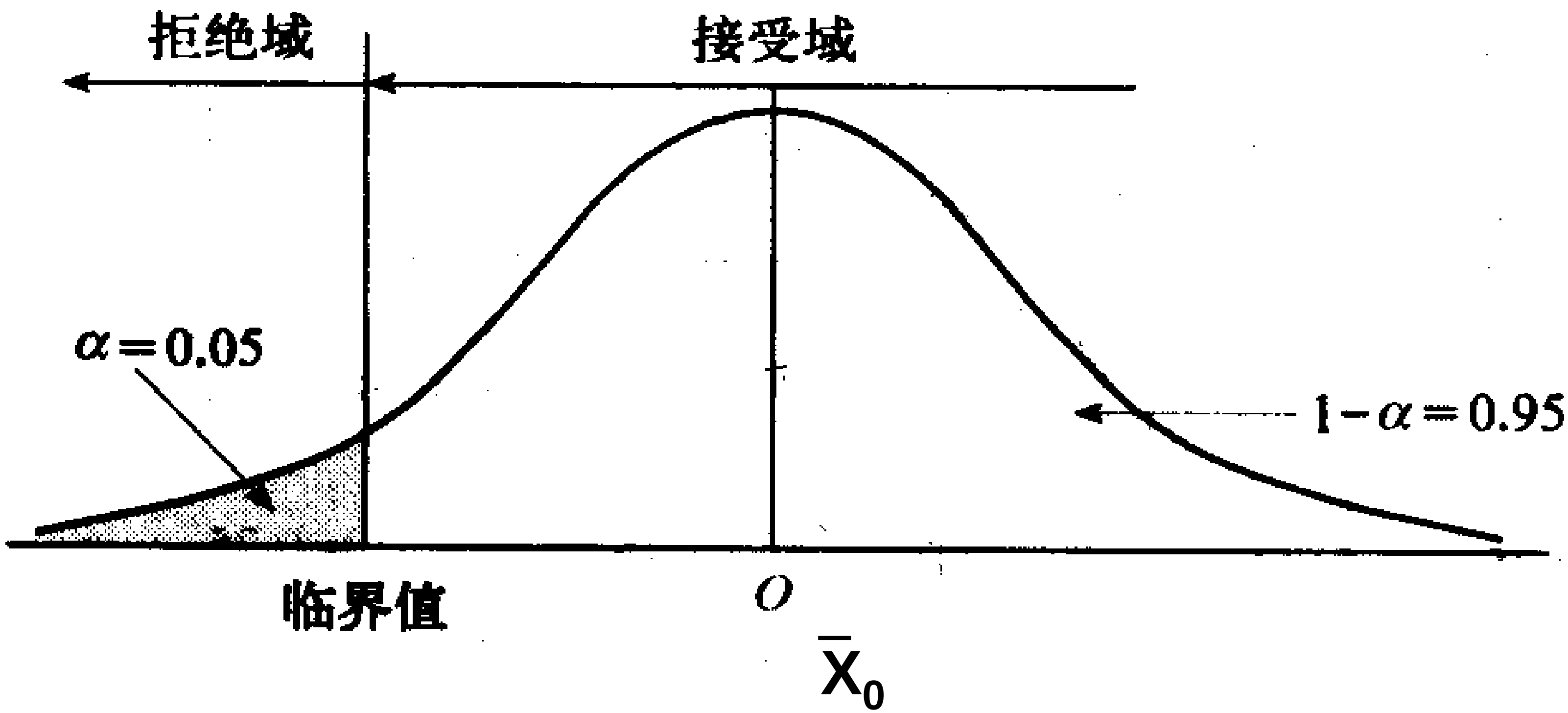
有一个临界值，一个拒绝域，拒绝域的面积为α。分为左侧检验和右侧检验两种情况。

单侧检验示意图（显著性水平与拒绝域）



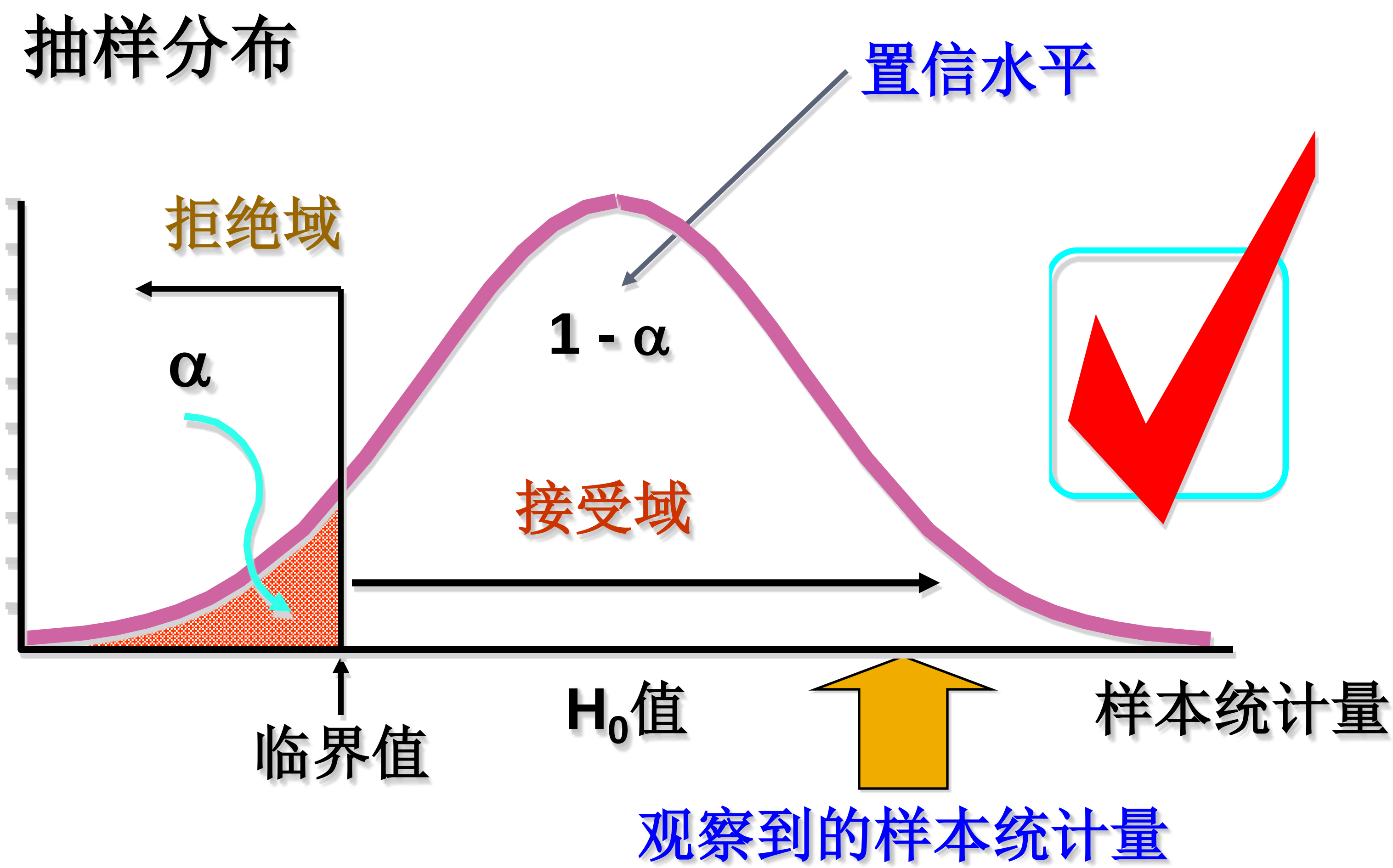
左侧检验

设H_0：$\bar{X} \geq \bar{X}_0$，H_1：$\bar{X} < \bar{X}_0$；临界值和拒绝域均在左侧。也称下限检验。

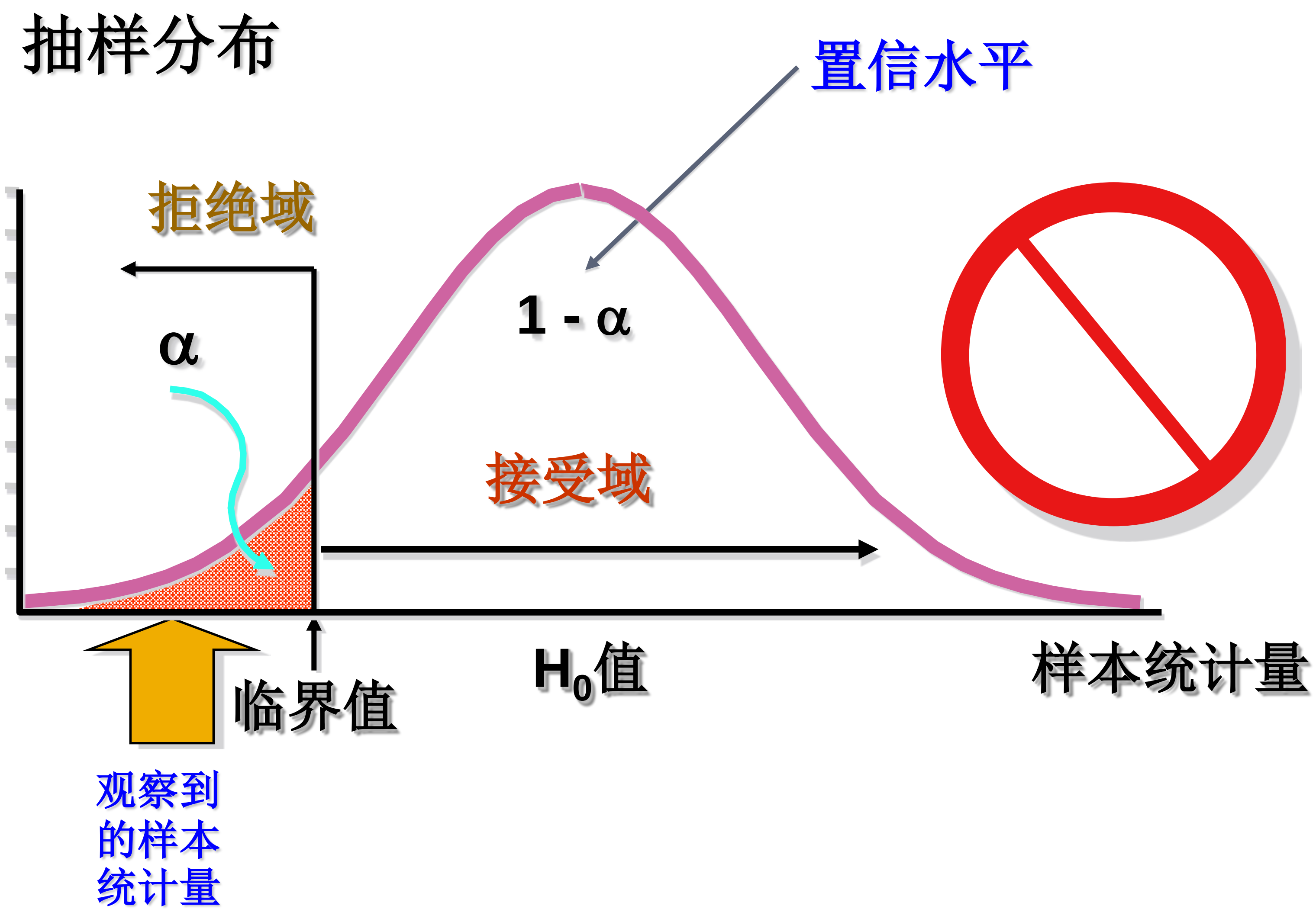


左单侧检验示意图

左侧检验示意图（显著性水平与拒绝域）

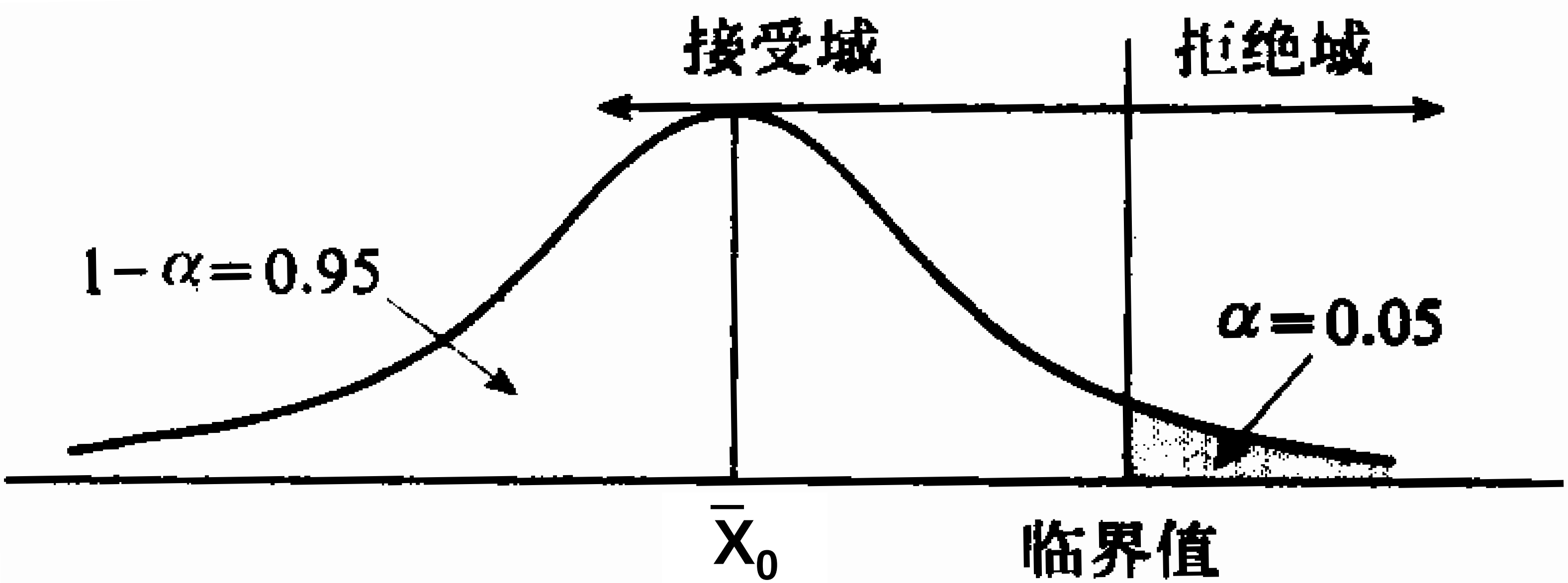


左侧检验示意图（显著性水平与拒绝域）



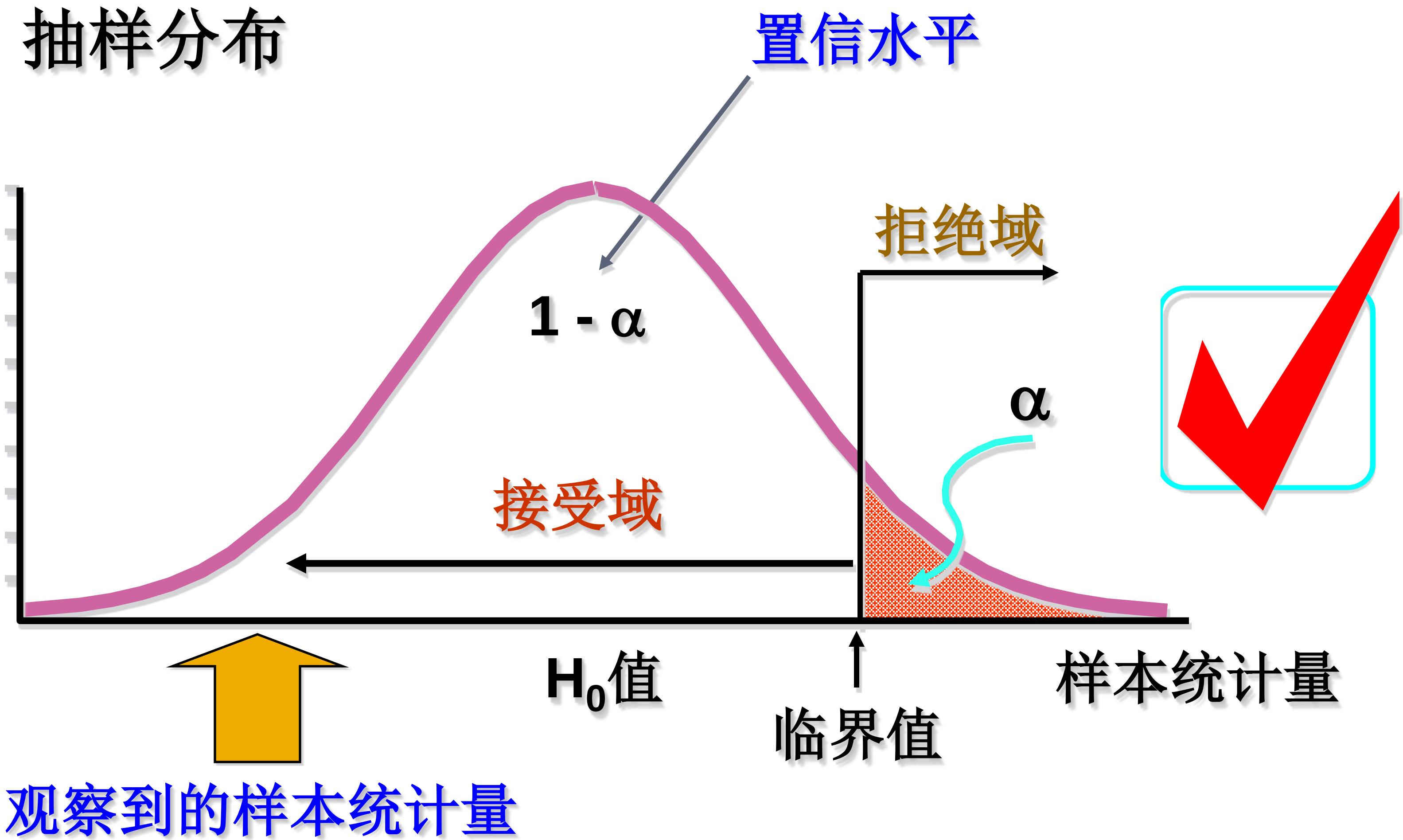
右侧检验

设H_0：$\bar{X} \le \bar{X}_0$，H_1：$\bar{X} > \bar{X}_0$；临界值和拒绝域均在右侧。也称上限检验。

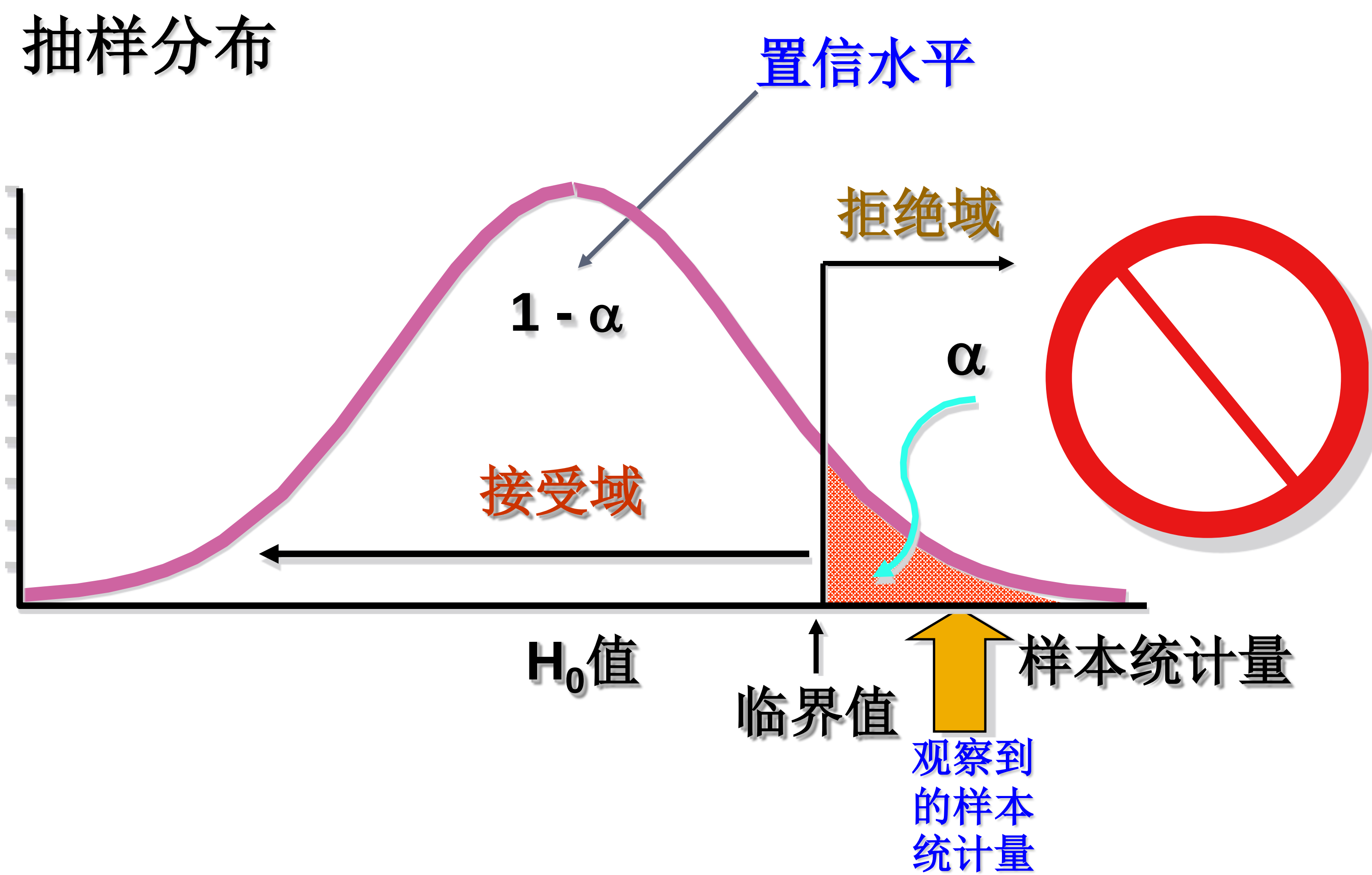


右单侧检验示意图

右侧检验示意图（显著性水平与拒绝域）



右侧检验示意图（显著性水平与拒绝域）



假设检验的两类错误

- 根据假设检验做出判断无非下述四种情况：

  1、原假设真实， 并接受原假设，判断正确；

  2、原假设不真实，且拒绝原假设，判断正确；

  3、原假设真实， 但拒绝原假设，判断错误；

  4、原假设不真实，却接受原假设，判断错误。
- 假设检验是依据样本提供的信息进行判断，有犯错误的可能。所犯错误有两种类型：
- 第一类错误是原假设H_0为真时，检验结果把它当成不真而拒绝了。犯这种错误的概率用α表示，也称作α型错误或弃真错误。
- 第二类错误是原假设H_0不为真时，检验结果把它当成真而接受了。犯这种错误的概率用β表示，也称作β型错误或取伪错误。

假设检验的两类错误

正确决策和犯错误的概率可以归纳为下表：

假设检验中各种可能结果的概率

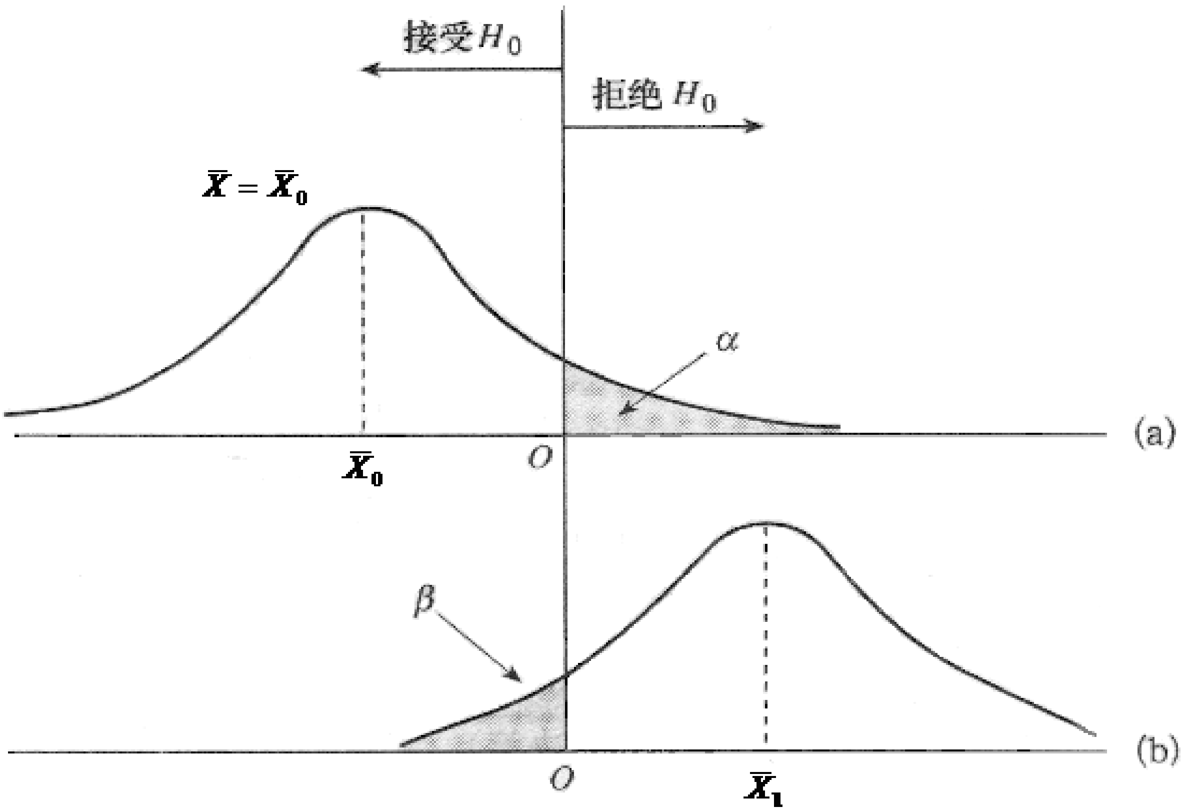
审判过程：H_0: 无罪

裁决	实际情况：无罪	实际情况：有罪
无罪	正确	错误
有罪	错误	正确

	接受H_0	拒绝H_0 接受H_1
H_0 为真	$1-\alpha$ (正确决策)	α(弃真错误)
H_0 为伪	β (取伪错误)	$1-\beta$(正确决策)

•假设检验两类错误关系的图示

以单侧上限检验为例，设H_0： $\bar{X} \leq \bar{X}_0$，H_1： $\bar{X} > \bar{X}_0$



图(a)
$\bar{X} \leqslant \bar{X}_0$
H_0为真

图(b)
$\bar{X} = \bar{X}_1 > \bar{X}_0$
H_0为伪

从上图可以看出，如果临界值沿水平方向右移，α将变小而β变大，即若减小α错误，就会增大犯β错误的机会；如果临界值沿水平方向左移，α将变大而β变小，即若减小β错误，也会增大犯α错误的机会。

两类错误的控制

⑴利用已知的实际总体参数值与假设参数值之间的大小关系，合理安排拒绝区域的位置

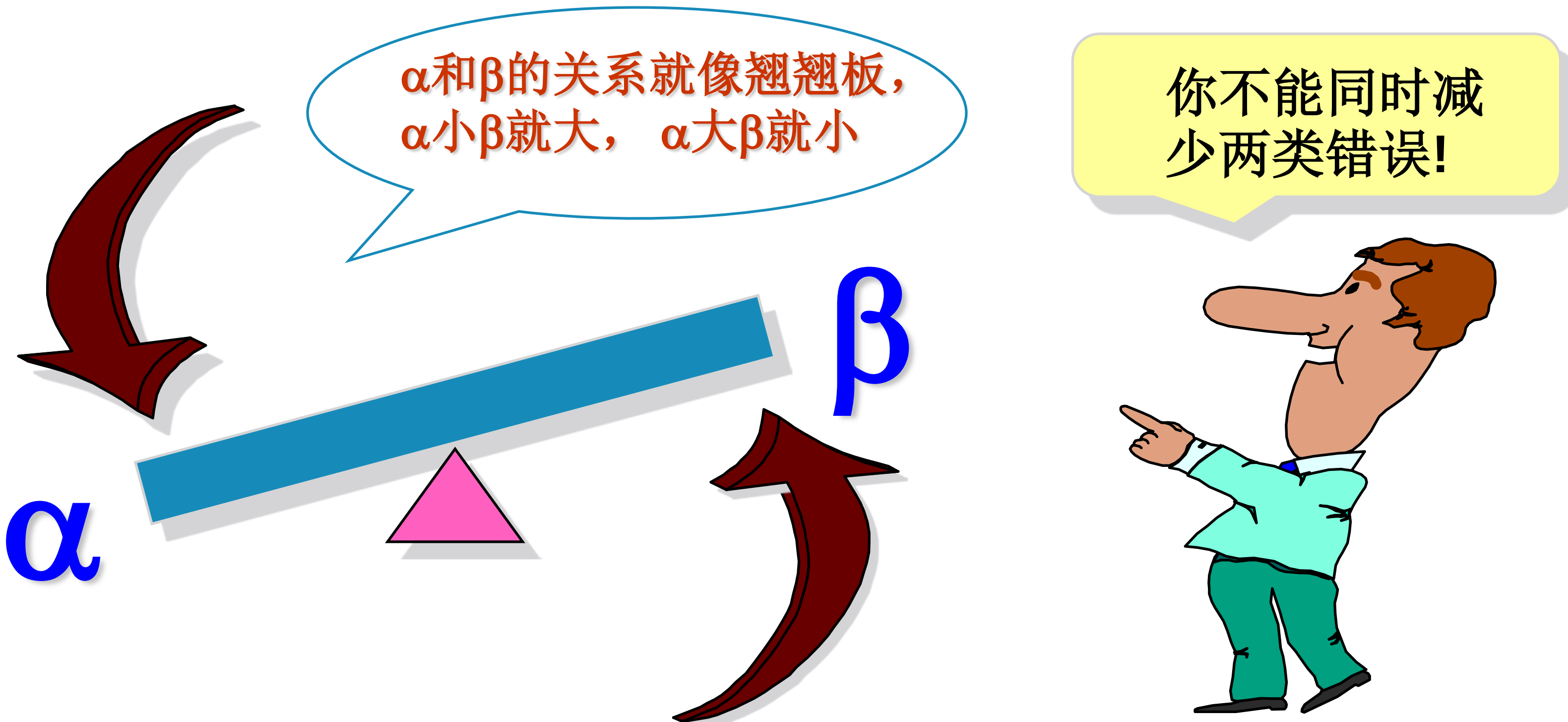
①若弃真的后果较大，则将α定小一点，β就大。

②若存伪的后果较大，则将β定小一点，α就大。

⑵使样本容量增大，可以同时减少两类错误，或减少其中一种错误而不致于增大另一种错误（如在α和其他条件不变时，β会减小）。（因为样本容量增大，抽样误差越小，样本分布就越高狭，两侧的面积就越小。）

α错误和β错误的关系

- 在样本容量n一定的情况下，假设检验不能同时做到犯α和β两类错误的概率都很小。若减小α错误，就会增大犯β错误的机会；若减小β错误，也会增大犯α错误的机会。要使α和β同时变小只有增大样本容量。但样本容量增加要受人力、经费、时间等很多因素的限制，无限制增加样本容量就会使抽样调查失去意义。因此假设检验需要慎重考虑对两类错误进行控制的问题。



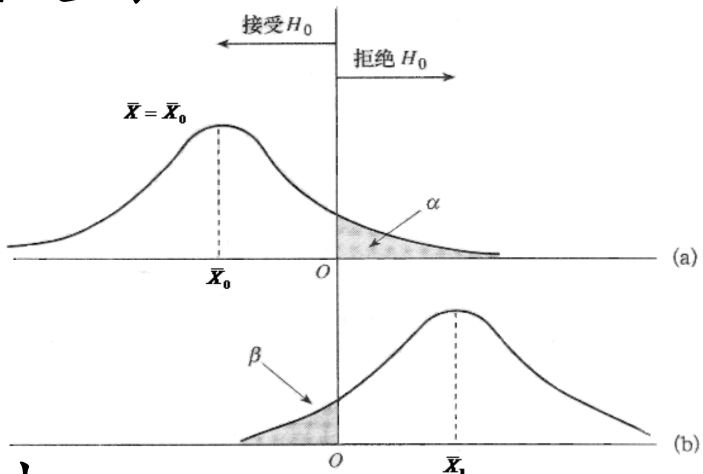
•两类错误的控制准则

- 假设检验中人们普遍执行同一准则：首先控制弃真错误（α错误）。假设检验的基本法则以α为显著性水平就体现了这一原则。
- 两个理由:
  - 统计推断中大家都遵循统一的准则，讨论问题会比较方便。
  - 更重要的是:原假设常常是明确的，而备择假设往往是模糊的。如H_0：$\bar{X}=\bar{X}_0$很清楚，而H_1：$\bar{X}\neq\bar{X}_0$则不太清楚，是$\bar{X}<\bar{X}_0$还是$\bar{X}>\bar{X}_0$？大多少小多少都不清楚。对含义清晰的数量标准进行检验更容易被接受。

因此，第一类错误成为控制两类错误的重点。

•两类错误的控制准则

- α错误：控制显著性水平。
  - 实验条件控制较好：α=0.05
  - 实验条件难于控制：α=0.01，或更高



P227 统计检验力 1- β

- β错误的影响因素与控制
  - 实际值与假设值相差越大，β越小。
  - α越小，β越大。同时控制，增加n。
  - α、n固定时，适当的检验类型可减小β。

• 假设检验的步骤

㈠根据研究需要提出原假设H_0和备择假设H_1

㈡确定适当的检验统计量

㈢确定显著性水平α和临界值及拒绝域

㈣根据样本数据计算检验统计量的值（或P值）

㈤将检验统计量值与临界值比较，作出拒绝或接受原假设的决策

假设检验的步骤
㈠根据研究需要提出原假设H_0和备择假设H_1

- 应该注意：

(1)对任一假设检验问题，其所有可能结果均应包括在所提出的两个对立假设中，原假设与对立假设总有一个、也只能有一个成立。

(2)原假设一定要有等号：＝或≤ 或 ≥。

- 原假设不是随意提出的，应该本着“不轻易拒绝原假设”的原则。

- 双侧检验原假设与备择假设的确定

- 双侧检验属于决策中的假设检验。即不论是拒绝H_0还是接受H_0，都必需采取相应的行动措施。
- 例如，某种零件的尺寸，要求其平均长度为10厘米，大于或小于10厘米均属于不合格。待检验问题是该企业生产的零件平均长度是10厘米吗?(属于决策中的假设)则建立的原假设与备择假设应为

$$H_0:\ \bar{X}=10 \qquad H_1:\ \bar{X}\neq 10$$

•单侧检验原假设与备择假设的确定

- 应区别不同情况采取不同的建立假设方法。

➡对于检验某项研究是否达到了预期效果

- 一般是将研究的预期效果（希望、想要证明的假设）作为备择假设H_1，将认为研究结果无效作为原假设H_0。先确立备择假设H_1。因为只有当检验结果与原假设有明显差别时才能拒绝原假设而接受备择假设，原假设不会轻易被拒绝，就使得希望得到的结论不会轻易被接受，从而减少结论错误。

❑ 例如，有研究预计，采用新技术生产后将会使某产品的使用寿命明显延长到1500小时以上。则建立的原假设与备择假设应为：　H_0: $\overline{X} \leq 1500$　H_1: $\overline{X} > 1500$

❑ 例如，有研究预计，改进生产工艺后会使某产品的废品率降低到2%以下。则建立的原假设与备择假设应为：

H_0: $\overline{X} \geq 2\%$　H_1: $\overline{X} < 2\%$

单侧检验原假设与备择假设的确定

- 对于检验某项**声明的有效性**

- 一般可将所作的声明作为原假设。将对该声明的质疑作为备择假设。先确立原假设H_0。因为除非有证据表明“声明”无效，否则就应认为该“声明”是有效的。

- 例如，某灯泡制造商声称，该企业生产的灯泡平均使用寿命在1000小时以上。通常除非样本能提供证据表明使用寿命在1000小时以下，否则就应认为厂商的声称是正确的。建立的原假设与备择假设应为：

$$H_0:\ \bar{X} \geq 1000 \qquad H_1:\ \bar{X} < 1000$$

- 对于上述问题还可以结合不同背景建立假设。同样的问题背景不同可以采用不同的原假设。
- 例如，一商店经常从某工厂购进某种商品，该商品质量指标为 $\bar{X}$， $\bar{X}$值愈大商品质量愈好。商店提出的进货条件是按批验收，只有通过假设“ $\bar{X} \geq \bar{X}_0$”检验的批次才能接受。有两种可能情况：

(1)如果根据过去较长时间购货记录，商店相信该厂产品质量好，于是同意把原假设定为 $\overline{X} \geqslant \overline{X}_0$，而且选择较低的检验显著性水平。这对工厂是有利的，使得达到质量标准的产品以很小的概率被拒收。虽然这会使商店面临接受不合标准产品的风险，但历史记录显示出现这种情况的可能性很小，而且商店也可因此获得较好的货源。

(2)如果过去一段时期的记录表明，该厂产品质量并不理想，商店则会坚持以 $\overline{X} \leqslant \overline{X}_0$为原假设，并选定较小的检验显著性水平。这对商店是有利的，不会轻易地拒绝原假设，有 $1-\alpha$的可能把劣质产品拒之门外。

双侧检验和单侧检验

1. **双侧检验：** 只强调差异而不强调方向性的检验称为双侧检验。
2. **单侧检验：** 强调某一方向的检验称为单侧检验。又分左侧检验和右侧检验。

通常适用于检验某一参数是否“大于”或“优于”、“快于”及“小于”、“劣于”、“慢于”另一参数等一类问题。

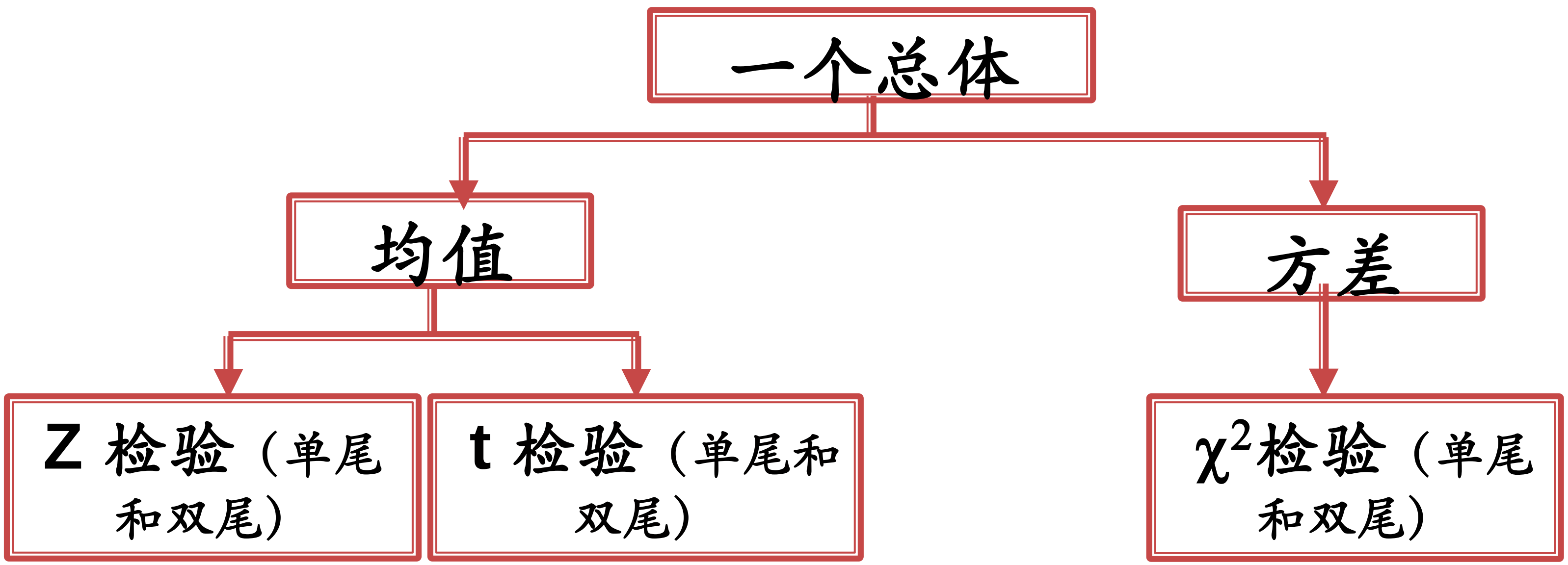
在实际研究中何时用单侧检验何时用双侧检验，一定要根据研究目的所规定的方向性来确定。

应该用单侧检验的问题，若使用双侧检验，其结果一方面可能使结论由“显著”变为“不显著”，另一方面也增大了β错误。

反之，若使用双侧检验的问题若用单侧来检验，虽然减小了β错误，但是使无方向性的问题人为地成为单方向问题，这也有悖于研究目的。

(二)确定适当的检验统计量

- 假设检验根据检验内容和条件不同需要采用不同的检验统计量。
- 在一个正态总体的参数检验中，Z统计量和t统计量常用于均值和比例的检验，χ^2统计量用于方差的检验。选择统计量需考虑的因素有被检验的参数类型、总体方差是否已知、用于检验的样本量大小等。



㈢确定显著性水平α和临界值及拒绝域

- 显著性水平α是当原假设为正确时被拒绝的概率，是由研究者事先确定的。
- 显著性水平的大小应根据研究需要的精确度和可靠性而定。通常取α＝0.05或α＝0.01，即接受原假设的决定是正确的可能性（概率）为95％或99％。
- 根据给定的显著性水平α，查表得出相应的临界值，同时指定拒绝域。

㈣根据样本数据计算检验统计量的值
㈤将检验统计量的值与临界值比较，作出拒绝或接受原假设的决策

- 例如，总体标准差σ已知时根据样本均值计算统计量Z的公式为

$$Z=\frac{\bar{x}-\bar{X}}{\sigma/\sqrt{n}}$$

- 如果检验统计量的值落入拒绝域，则拒绝原假设，接受备择假设；如果检验统计量的值落入接受域，则接受原假设，拒绝备择假设。

第二/三节 平均数差异显著性检验

一、平均数差异显著性检验的类型

1. 单总体平均数差异显著性检验，也叫平均数的显著性检验。
2. 双总体平均数差异显著性检验，也叫平均数差异的显著性检验。

二、平均数差异显著性检验的步骤：

1. 建立假设：

双侧检验的假设：$H_0: \mu = \mu_{0,} H_1: \mu \neq \mu_0$

或者(双总体)$H_0: \mu_1 = \mu_2, H_1: \mu_1 \neq \mu_2$

单侧检验的假设：$H_0: \mu \leq \mu_0, H_1: \mu > \mu_0$

或者(双总体)$H_0: \mu_1 \leq \mu_2, H_1: \mu_1 > \mu_2$

2. 根据给定条件确定抽样分布的形态，确定相应的检验方法，并计算出统计量的值。
3. 确定α，查出理论值，从而确定出H_0拒绝和接受区域。
4. 作出判断。如$Z>Z_{临}$,即$p<\alpha$,则拒绝H_0,接受H_1。

三、单总体平均数差异显著性检验

㈠总体分布为正态分布,且总体方差已知时

$$Z=\frac{\overline{X}-\mu_0}{SE_{\overline{X}}}\cdots(8-1),\text{其中}SE_{\overline{X}}=\frac{\sigma_0}{\sqrt{n}}$$

例1 某地区统考数学,假设该统考数学成绩服从正态分布,已知其总平均分为50分,标准差为12分。从该地区随机选择一个班作为样本,该班有学生50人,经计算该班平均成绩为53分,试问该班成绩与总成绩的差异是否显著。

例1的计算

解：①建立假设：$H_0: \mu = \mu_0, H_1: \mu \neq \mu_0$

②计算统计量：$SE_{\overline{X}} = \dfrac{\sigma_0}{\sqrt{n}} = \dfrac{12}{\sqrt{50}} = \dfrac{12}{7.07} = 1.7$

$$Z = \frac{\overline{X} - \mu_0}{SE_{\overline{X}}} = \frac{53 - 50}{1.7} = 1.76$$

③判断结果:因为1.76<1.96,所以P>0.05

差异不显著,接受H_0,拒绝H_1。

故该班数学成绩与总平均成绩的差异不显著。

单侧检验的例题

例2 对某专业课在全国同类高校内进行统一测试，已知全体考生成绩服从正态分布，其总平均分为64分，总标准差为8.6分，从某高校随机抽取20份试卷，经计算得到这20份试卷的平均成绩为70分，问该校学生的平均成绩是否 显著地优于全体学生的平均水平?

例2的计算

解：①建立假设：$H_0: \mu \le \mu_0, H_1: \mu > \mu_0$

②计算统计量：

$$SE_{\overline{X}} = \frac{\sigma_0}{\sqrt{n}} = \frac{8.6}{\sqrt{20}} = \frac{8.6}{4.47} = 1.92$$

$$Z = \frac{\overline{X} - \mu_0}{SE} = \frac{70-64}{1.92} = 3.125$$

③查正态表（单侧）得：

当 $\alpha = 0.01$时，$Z_{0.01} = 2.33$

④判断结果:因为3.125>2.33,所以P<0.01。差异极其显著，拒绝H₀。故该校学生的平均成绩极显著地优于全体学生的平均水平。

三、单总体平均数差异显著性检验

㈡总体分布为正态,总体方差未知

当总体分布为正态分布,总体方差未知时,样本平均数的抽样分布服从自由度为n-1 的t分布,此时可用t检验方法进行检验。在检验中，计算统计量的值的公式为:

$$t=\frac{\overline{X}-\mu_0}{SE_{\overline{X}}}\ldots(8-2),其中SE_{\overline{X}}=\frac{S_n}{\sqrt{n-1}}=\frac{S_{n-1}}{\sqrt{n}}$$

例3 已知4岁正常男童平均体重为15.6公斤,从某幼儿园中随机抽取20名4岁男童,经测量计算出这20名男童平均体重为16公斤,标准差为1.76公斤,试问该幼儿园4岁男童的体重与4岁正常男童平均体重有无显著性差异?

例3的计算

解:①建立假设: $H_0: \mu = \mu_0, H_1: \mu \neq \mu_0$

②计算统计量:

$$SE_{\overline{X}} = \frac{S_n}{\sqrt{n-1}} = \frac{1.76}{\sqrt{20-1}} = 0.4$$

$$t = \frac{\overline{X} - \mu_0}{SE_{\overline{X}}} = \frac{16-15.6}{0.4} = 1$$

③判断结果: 当$\alpha = 0.05$时, $df = 20-1 = 19$时,

$$t_{(19)0.05/2} = 2.093$$

因为1<2.093,所以,P>0.05，差异不显著。

故该幼儿园4岁男童平均体重与正常男童平均体重无显著差异。

例题

例4 某实验组随机选择了40名儿童作提高儿童智力水平实验，实验结束后，对参加实验的儿童进行韦氏儿童智力测验，结果得到这40名儿童的智商平均值为105，标准差为12，已知韦氏儿童智力测验的平均值为100，试问该实验是否成功？

解：①建立假设：$H_0:\mu \le \mu_0, H_1:\mu > \mu_0$

②计算统计量：$SE_{\overline{X}} = \dfrac{S}{\sqrt{n-1}} = \dfrac{12}{\sqrt{40-1}} = \dfrac{12}{6.24} = 1.92$

$$t = \frac{\overline{X}-\mu_0}{SE_{\overline{X}}} = \frac{105-100}{1.92} = 2.6$$

例4的计算（续）

③判断结果: 当$\alpha = 0.01, df = 39$时, $t_{(39)0.01} = 2.423$

因为$2.6 > 2.423$, 所以$P < 0.01$

故实验后的儿童智力水平极明显高于同龄儿童一般水平,实验取得了成功。

三、单总体平均数差异显著性检验

㈢总体为非正态分布

当总体为非正态分布时，根据中心极限定理，样本容量足够大时（n≥30，也有认为n≥50)，样本平均数的抽样分布趋于正态分布，在平均数差异显著性检验中，也可近似地采用Z检验方法进行检验。

当总体方差已知时,其检验公式为:

$$Z' = \frac{\overline{X} - \mu_0}{SE_{\overline{X}}} \qquad \text{其中} \qquad SE_{\overline{X}} = \frac{\sigma_0}{\sqrt{n}}$$

当总体方差未知时，上式中的σ_0可用它的无偏估计量S_{n-1}代替，

当样本量足够大时，也可以用S代替σ。

$$SE_{\overline{X}} = \frac{S_{n-1}}{\sqrt{n}}$$

注意:这种采用近似Z检验的方法只能适用n>30时的情况,当n<30时不适用这种方法,也不能用t检验方法,只能采用非参数检验方法进行检验。

例5 某县进行物理竞赛，结果分数的分布不是正态，总平均成绩为55分。其中某重点学校参加竞赛60人，其平均成绩为62分，标准差为14分，试问该校平均成绩是否明显优于全县一般水平。

四、双总体平均数差异显著性检验

当两个变量间的相关系数为r时，两变量的方差为：(P118)

$$\sigma^2_{(X-Y)} = \sigma^2_X + \sigma^2_Y - 2r\sigma_x\sigma_y$$

当两个变量相互独立时，其和（或差）的方差等于各自方差的和：

$$\sigma^2_{(X\pm Y)} = \sigma^2_X + \sigma^2_Y$$

因此有：

$$SE_{D_{\overline{X}}} = \sqrt{SE_1^2 + SE_2^2} = \sqrt{\frac{\sigma_1^2}{n_1} + \frac{\sigma_2^2}{n_2} - 2r\frac{\sigma_1}{\sqrt{n_1}}\frac{\sigma_2}{\sqrt{n_2}}}$$

当两个样本独立时，r=0，这时上述公式为

$$SE_{D_{\overline{X}}} = \sqrt{SE_1^2 + SE_2^2} = \sqrt{\frac{\sigma_1^2}{n_1} + \frac{\sigma_2^2}{n_2}}$$

四、双总体平均数差异显著性检验

4.1 两总体分布为正态,且方差已知当两样本相互独立或相关

两平均数之差的标准误为:

$$SE_{D_{\overline{X}}} = \frac{S_{D_{\overline{X}}}}{\sqrt{n}} = \sqrt{\frac{\sigma_1^2 + \sigma_2^2 - 2r\sigma_1\sigma_2}{n}}$$

$$= \sqrt{SE_1^2 + SE_2^2 - 2rSE_1SE_2} \ldots (8-4)$$

检验统计量公式为

$$Z = \frac{(X_1 - X_2) - (\mu_1 - \mu_2)}{SE_{D_{\overline{X}}}} = \frac{D_{\overline{X}} - \mu_{D_{\overline{X}}}}{SE_{D_{\overline{X}}}} \qquad \cdots (8-5)$$

四、双总体平均数差异显著性检验

4.1 两总体分布为正态,且方差已知

(4.1.1) 两样本相互独立

例6: 从某科高考成绩中,从甲省抽取80名考生,计算平均成绩为84分,从乙省抽取100名考生,计算平均成绩为80分,已知甲乙两省该科高考成绩均呈正态分布,甲省标准差为12.4分,乙省标准差为11.2分,试问甲乙两省该科高考成绩是否有显著性差异?

例6计算

解： 1. 建立假设： $H_0: \mu_1 = \mu_2, H_1: \mu_1 \neq \mu_2$

2. 计算检验的统计量：

$$SE_{D_{\overline{X}}} = \sqrt{\frac{\sigma_1^2}{n_1} + \frac{\sigma_2^2}{n_2}} = \sqrt{\frac{12.4^2}{80} + \frac{11.2^2}{100}}$$

$$= \sqrt{1.922 + 1.2544} = 1.7822$$

$$Z = \frac{\overline{X}_1 - \overline{X}_2}{SE_{D_{\overline{X}}}} = \frac{84 - 80}{1.7822} = 2.2444$$

3. 判断结果:查表得， $Z_{0.05/2} = 1.96$ 因为2.24>1.96,所以P<0.05,差异极其显著。

故根据这次抽样，可以推断出甲乙两省该科高考成绩有显著性差异。

例题

例7从某市城区随机抽取5岁男童30名,测得其平均体重为20.15公斤,从该市郊区随机抽取25名5岁男童,测得其平均体重为18.25公斤,已知该市城区5岁男童体重的标准差为1.85公斤,郊区5岁男童体重的标准差为1.78公斤.能否根据这次抽样推断出该市城区与郊区5岁男童平均体重的差异是否显著。

例7计算

解： 1. 建立假设: $H_0: \mu_1 = \mu_2, H_1: \mu_1 \neq \mu_2$

2. 计算检验的统计量:

$$SE_{D_{\overline{X}}} = \sqrt{\frac{\sigma_1^2}{n_1} + \frac{\sigma_2^2}{n_2}} = \sqrt{\frac{1.85^2}{30} + \frac{1.78^2}{25}}$$

$$= \sqrt{0.114 + 0.127} = 0.49$$

$$Z = \frac{\overline{X}_1 - \overline{X}_2}{SE_{D_{\overline{X}}}} = \frac{20.15 - 18.25}{0.49} = 3.87$$

3. 判断结果:查表得, $Z_{0.01/2} = 2.58$ 因为3.87>2.58,所以 P<0.01,差异极其显著。

故根据这次抽样，可以推断出该市郊区5岁男童平均体重有极其显著的差异。

四、双总体平均数差异显著性检验

4.1 两总体分布为正态,且方差已知

(4.1.2) 两样本相关

例8 某实验组选择20名儿童作智力发展实验在实验前和实验后对实验儿童进行一次韦氏儿童智力测验，在实验前得到平均智商为102，实验后所得平均智商为107，已知韦氏儿童智力测验标准差为15，两次测验相关系数为0.7，试问该实验是否取得了成功。

例8 计算

解： 1. 建立假设： $H_0:\mu_1 \ge \mu_2, H_1:\mu_1 \prec \mu_2$

2. 计算检验的统计量:

$$SE_{D_{\overline{X}}} = \sqrt{\frac{\sigma_1^2+\sigma_2^2-2r\sigma_1\sigma_2}{n}}$$

$$= \sqrt{\frac{15^2+15^2-2\times0.7\times15\times15}{20}} = 2.6$$

$$Z = \frac{\overline{X}_1-\overline{X}_2}{SE_{D\overline{X}}} = \frac{102-107}{2.6} = -1.92$$

3. 判断结果:

查表得: $Z_{0.05} = 1.645, Z_{0.01} = 2.33$

因为 $1.645 < |-1.92| < 2.33$, 所以, $0.05 > P > 0.01$,

例题

例9 某地两年内进行了物理竞赛,已知两次竞赛成绩均服从正态分布, 两次竞赛的标准差分别为12分和14分,某校有10 名学生先后两次参加了竞赛,根据他们的竞赛成绩计算出他们的平均成绩分别为72分和70分,这10 名学生竞赛成绩的相关系数为0.76,试问该地两次竞赛成绩有无显著性差异。

例9 计算

解： 1. 建立假设： $H_0: \mu_1 = \mu_2, H_1: \mu_1 \neq \mu_2$

2. 计算检验的统计量:

$$SE_{D_{\overline{X}}} = \sqrt{\frac{\sigma_1^2 + \sigma_2^2 - 2r\sigma_1\sigma_2}{n}}$$

$$= \sqrt{\frac{12^2 + 14^2 - 2 \times 0.76 \times 12 \times 14}{10}} = 2.91$$

$$Z = \frac{\overline{X_1} - \overline{X_2}}{SE_{D\overline{X}}} = \frac{72 - 70}{2.91} = 0.6875$$

3. 判断结果:

判断结果:查表得,$Z_{0.05/2} = 1.96$ 因为0.6875<1.96,所以P>0.05,差异不显著。

四、双总体平均数差异显著性检验

4.2 两总体都是正态分布,且方差未知

当两总体为正态分布,两总体方差均未知,从两总体所抽出两样本平均数的差数服从df=n-1的t分布,可用t检验方法进行检验。

4.2.1 两样本容量不等，且为独立样本

根据公式(8-5)则标准误的计算公式为：$SE_{D_{\overline{X}}} = \sqrt{\dfrac{\sigma_1^2}{n_1} + \dfrac{\sigma_2^2}{n_2}}$

在该式中，$\sigma 1^2$与$\sigma 2^2$都未知，

此时可用无偏方差来代替

$$SE_{D_{\overline{X}}} = \sqrt{\frac{S_{n_1-1}^2}{n_1} + \frac{S_{n_2-1}^2}{n_2}} \ldots (8-7)$$

四、双总体平均数差异显著性检验

4.2 两总体都是正态分布,且方差未知

4.2.1 两样本容量不等，且为独立样本

第一，当总体方差相等时，

其检验公式为：
$$t=\frac{D_{\overline{X}}-\mu_{D_{\overline{X}}}}{SE_{D_{\overline{X}}}}=\frac{\overline{X}_1-\overline{X}_2}{SE_{D_{\overline{X}}}}\cdots(8-8)$$

$df=n_1+n_2-2$

其中

$$SE_{D_{\overline{X}}}=\sqrt{\frac{n_1S_1^2+n_2S_2^2}{n_1+n_2-2}\cdot\frac{n_1+n_2}{n_1n_2}}$$

当总体方差相等时，方差的估计

联合方差$s_p{}^2$

$$s_p^2=\frac{(n_1-1)s_{n_1-1}^2+(n_2-1)s_{n_2-1}^2}{(n_1-1)+(n_2-1)}=\frac{n_1s_{n_1}^2+n_2s_{n_2}^2}{(n_1-1)+(n_2-1)}$$

$$SE_{D_{\overline{X}}}=\sqrt{s_p^2\left(\frac{1}{n_1}+\frac{1}{n_2}\right)}=\sqrt{\frac{n_1S_1^2+n_2S_2^2}{n_1+n_2-2}\cdot\frac{n_1+n_2}{n_1n_2}}$$

精确合成参见P90 公式4-13

例题

例10 某校进行速度测验，共选择40名参加，其中男生22人，女生18人，男生测验的平均分为72分，标准差为8分，女生测验的平均分为70分，标准差为7.6分。已知该速度测验结果服从正态分布，试问男女生速度测验的平均结果有无显著性差异。

例10的计算
(为了简单化，本题暂先假设检验得出方差相等)

解：①建立假设：　$H_0: \mu_1 = \mu_2, H_1: \mu_1 \neq \mu_2$

②计算检验的统计量:

$$SE_{D_{\overline{X}}} = \sqrt{\frac{n_1 S_1^2 + n_2 S_2^2}{n_1 + n_2 - 2} \cdot \frac{n_1 + n_2}{n_1 n_2}}$$

$$= \sqrt{\frac{22 \times 8^2 + 18 \times 7.6^2}{22 + 18 - 2} \times \frac{22 + 18}{22 \times 18}} = 2.55$$

$$t = \frac{\overline{X_1} - \overline{X_2}}{SE_{D_{\overline{X}}}} = \frac{72 - 70}{2.55} = 0.78$$

③判断结果：当α=0.05，df=n1+n2-2 = 22+18-2=38时，$t_{(38)0.05/2}$=2.021。

因为0.78<2.021,所以P>0.05,差异不显著。

四、双总体平均数差异显著性检验

4.2 两总体都是正态分布,且方差未知

4.2.1 两样本容量不等，且为独立样本

第二，当总体方差不等时，

此时的样本平均数之差不再服从t分布，不能再运用t检验方法来检验了。这时常用柯克兰—柯克斯检验方法。

标准误的计算公式为：

$$SE_{D_{\overline{X}}} = \sqrt{\frac{S_{n_1-1}^2}{n_1} + \frac{S_{n_2-1}^2}{n_2}} = \sqrt{\frac{S_1^2}{n_1-1} + \frac{S_2^2}{n_2-1}} \ldots (8-11)$$

四、双总体平均数差异显著性检验

4.2 两总体都是正态分布,且方差未知

4.2.1 两样本容量不等，且为独立样本

第二，当总体方差不等时，

检验的统计量公式为：

$$t' = \frac{D_{\overline{X}} - \mu_{D_{\overline{X}}}}{SE_{D_{\overline{X}}}} = \frac{\overline{X_1} - \overline{X_2}}{SE_{D_{\overline{X}}}} \ldots (8-12)$$

t'的分布理论值为：

$$t'_{(\alpha)} = \frac{SE^2_{\overline{X}_1} \cdot t_{1(\alpha)} + SE^2_{\overline{X}_2} \cdot t_{2(\alpha)}}{SE^2_{\overline{X}_1} + SE^2_{\overline{X}_2}} \ldots (8-13)$$

$t_1(\alpha)$ 为样本1的自由度df1=n1-1对应的临界值

$t_2(\alpha)$ 为样本1的自由度df2=n2-1对应的临界值

例11 用例10的数据，假如方差不齐性，试求男女生测验成绩差异是否 显著。

- 例10 某校进行速度测验，共选择40名参加，其中男生22人，女生18人，男生测验的平均分为72分，标准差为8分，女生测验的平均分为70 分，标准差为7.6分。已知该速度测验结果服从正态分布，试问男女生速度测验的平均结果有无显著性差异。

例11 计算

解：①建立假设：$H_0: \mu_1 = \mu_2, H_1: \mu_1 \neq \mu_2$

②计算检验统计量：

$$SE_{D_{\overline{X}}} = \sqrt{\frac{8^2}{22-1} + \frac{7.6^2}{18-1}} = 2.54$$

$$t' = \frac{\overline{X_1} - \overline{X_2}}{SE_{D_{\overline{X}}}} = \frac{72-70}{2.54} = 0.79$$

$$SE_{\overline{X}_1} = \frac{8}{\sqrt{22-1}} = 1.75, SE_{\overline{X}_2} = \frac{7.6}{\sqrt{18-1}} = 1.84$$

$$t_{0.05/2} = \frac{1.75^2 \times 2.08 + 1.84^2 \times 2.11}{1.75^2 + 1.84^2} = 2.09$$

③判断结果：因为0.79<2.09,所以P>0.05,差异不显著。

四、双总体平均数差异显著性检验

4.2 两总体都是正态分布,且方差未知

4.2.2 两样本容量相等

①独立样本:

a. 方差不等

其标准误计算公式为: $SE_{D_{\overline{X}}} = \sqrt{\dfrac{S_{n_1-1}^2 + S_{n_2-1}^2}{n}} = \sqrt{\dfrac{S_1^2 + S_2^2}{n-1}}$

其检验公式还是:

$$t' = \frac{D_{\overline{X}} - \mu_{D_{\overline{X}}}}{SE_{D_{\overline{X}}}} = \frac{\overline{X}_1 - \overline{X}_2}{SE_{D_{\overline{X}}}} \ldots (8-12)$$

$$t'_{(\alpha)} = \frac{SE_{\overline{X}_1}^2 \cdot t_{1(\alpha)} + SE_{\overline{X}_2}^2 \cdot t_{2(\alpha)}}{SE_{\overline{X}_1}^2 + SE_{\overline{X}_2}^2} = t_{1(\alpha)} = t_{2(\alpha)}$$

$df = n$-1

四、双总体平均数差异显著性检验

4.2 两总体都是正态分布,且方差未知

4.2.2 两样本容量相等

①独立样本:

a. 方差相等

其检验公式还是(与样本量无关):

$$t=\frac{D_{\overline{X}}-\mu_{D_{\overline{X}}}}{SE_{D_{\overline{X}}}}=\frac{\overline{X}_1-\overline{X}_2}{SE_{D_{\overline{X}}}}\ldots(8-8)$$

其标准误计算公式为:

$$SE_{D_{\overline{X}}}=\sqrt{\frac{S_{n_1-1}^2+S_{n_2-1}^2}{n}}=\sqrt{\frac{S_1^2+S_2^2}{n-1}}$$

$df = 2n-2$

$$s_p^2=\frac{(n_1-1)s_{n_1-1}^2+(n_2-1)s_{n_2-1}^2}{(n_1-1)+(n_2-1)}=\frac{\left(s_{n_1-1}^2+s_{n_2-1}^2\right)}{2}=\frac{n_1 s_{n_1}^2+n_2 s_{n_2}^2}{(n_1-1)+(n_2-1)}=\frac{n\left(s_{n_1}^2+s_{n_2}^2\right)}{2(n-1)}$$

$$SE_{D_{\overline{X}}}=\sqrt{s_p^2\left(\frac{1}{n}+\frac{1}{n}\right)}=\sqrt{\frac{S_{n_1-1}^2+S_{n_2-1}^2}{2}\cdot\frac{2}{n}}=\sqrt{\frac{S_{n_1-1}^2+S_{n_2-1}^2}{n}}=\sqrt{\frac{n\left(S_{n_1}^2+S_{n_2}^2\right)}{2(n-1)}\cdot\frac{2}{n}}=\sqrt{\frac{S_{n_1}^2+S_{n_2}^2}{n-1}}$$

例题

例12 某市进行一次物理统考, 已知统考成绩服从正态分布,从该市A校和B校各抽取20名学生,经计算两校的物理统考平均成绩分别为82分和77分,标准差分别为7.2分和7.5分,试问A、B两校在该次物理统考中的平均成绩差异是否显著。

例12的计算

解：①建立假设:　$H_0:\mu_1=\mu_2, H_1:\mu_1\neq\mu_2$

②计算检验统计量:

$$SE_{D_{\overline{X}}}=\sqrt{\frac{7.2^2+7.5^2}{20-1}}=2.38$$

$$t=\frac{82-77}{2.38}=2.10$$

③判断结果:查表得:　$t_{(38)0.05/2}=2.024, t_{(38)0.01/2}=2.712$

因为2.024<2.10<2.712,所以0.05>P>0.01, α=0.05时差异显著。故该市AB两校在该次物理统考中的平均成绩差异显著。

四、双总体平均数差异显著性检验

4.2 两总体都是正态分布,且方差未知

4.2.2 两样本(X_{1i}, X_{2i})容量相等

②相关样本（t检验一般不需要事先进行方差齐性检验，见P242说明）

令$d_i=X_{1i}-X_{2i}$为每一对数据之差，则对$\bar{d}$的显著性检验就等同于对两样本均值差异的显著性检验：

T检验

$$t=\frac{\bar{X}_1-\bar{X}_2}{SE_{D_{\bar{X}}}}\ldots(8-17)$$

自由度为 $df=n-1$

其标准误计算公式为:

1. 相关系数未知，

$$SE_{D_{\bar{x}}}=SE_{\bar{d}}=\sqrt{\frac{s_d^2}{n-1}}=\sqrt{\frac{\sum d^2-\frac{\left(\sum d\right)^2}{n}}{n(n-1)}}$$

P88 s_d公式4-9

2. 相关系数已知，

$$SE_{D_{\bar{x}}}=\sqrt{\frac{s_1^2+s_2^2-2rs_1^2s_2^2}{n-1}}$$

例题

例13 某实验小组进行一项学习方法实验,在实验前,经过精心匹配,分成两小组,一为实验组,一为对照组,各组均为20人。实验结束后，对两组均进行智力测验,经计算，其智商成绩分别为：实验组平均智商为106，标准差为16；对照组平均智商为104，标准差为15，两组测验的相关系数为0.75,试问该学习方法对提高智力是否有显著影响。

例13的计算

解：①建立假设: $H_0: \mu_1 \le \mu_2, H_1: \mu_1 > \mu_2$(单侧检验)

②计算统计量:

$$SE_{D_{\overline{X}}} = \sqrt{\frac{16^2 + 15^2 - 2 \times 0.75 \times 16 \times 15}{20-1}} = 2.52$$

$$t = \frac{106-104}{2.52} = 0.79$$

③判断结果:查表得: $t_{(19)0.05} = 2.093$

因为0.79<2.093,所以P>0.05,差异不显著。

故该学习方法实验对学生智力提高没有显著影响。

例题8-10,8-11

例8-10 对9个被试进行两种夹角（15°,30°）的缪勒-莱依尔错觉实验结果如下，问两种夹角的情况下错觉量是否显著差异。

例8-11 用相关系数已知的公式来计算本题。

被试	1	2	3	4	5	6	7	8	9
15°	14.7	18.9	17.2	15.4	15.3	13.9	20.0	16.2	15.3
30°	10.6	15.1	16.2	11.2	12.0	14.7	18.1	13.8	10.9
d_i									

解：

8-10：先计算d_i，$\overline{X}_1$，$\overline{X}_2$，接着计算s_d，然后计算t统计量(t1=4.6248)

8-11：计算积差相关 0.7451 (t2=4.6248)，Spearman相关0.7029 (t3=4.32)

$$t_1=t_2$$

练习题（首师大2003研，浙大2006研）

- 随机从某总体选取10名被试，分别实施两次数学测验，两次测验的成绩见下表，问被试在两次测验的平均数是否有显著差异？试对结果进行解释（$\alpha=0.05$，$df=9$，$t=2.262$；$df=18$，$t=2.552$）

表　　10 名被试两次测验的成绩

被试	1	2	3	4	5	6	7	8	9	10
测验一	65	48	63	52	61	53	63	70	65	66
测验二	61	42	66	52	47	58	65	62	64	69

四、双总体平均数差异显著性检验

4.3 两总体为非正态分布

当样本容量都大于30(或都大于50时)

4.3.1 独立样本

①当总体方差已知时,其检验公式为:

$$Z' = \frac{D_{\overline{X}} - \mu_{D_{\overline{X}}}}{\sqrt{\frac{\sigma_1^2}{n_1} + \frac{\sigma_2^2}{n_2}}} \cdots (8-19)$$

②当总体方差未知时,其检验公式为:

$$Z' = \frac{D_{\overline{X}} - \mu_{D_{\overline{X}}}}{\sqrt{\frac{S_{n_1-1}^2}{n_1} + \frac{S_{n_2-1}^2}{n_2}}} = \frac{D_{\overline{X}} - \mu_{D_{\overline{X}}}}{\sqrt{\frac{S_1^2}{n_1-1} + \frac{S_2^2}{n_2-1}}} \cdots (8-20)$$

四、双总体平均数差异显著性检验

4.3 两总体为非正态分布

当样本容量都大于30(或都大于50时)

4.3.1 相关样本:

①当总体方差已知时,其检验公式为:

$$Z' = \frac{\overline{X}_1 - \overline{X}_2}{\sqrt{\frac{\sigma_1^2 + \sigma_2^2 - 2r\sigma_1\sigma_2}{n}}} \cdots (8-21)$$

②当总体方差未知时,其检验公式为:

$$Z' = \frac{\overline{X}_1 - \overline{X}_2}{\sqrt{\frac{S_{n_1-1}^2 + S_{n_2-1}^2 - 2rS_{n_1-1}S_{n_2-1}}{n}}} = \frac{\overline{X}_1 - \overline{X}_2}{\sqrt{\frac{S_1^2 + S_2^2 - 2rS_1S_2}{n-1}}} \cdots (8-22)$$

平均数差异的显著性检验小结

总体分布	总体方差		样本情况		检验方法
正态分布	已知		独立		Z
			相关		Z
	未知	$\sigma_1^2=\sigma_2^2$	独立		t（$df=n_1+n_2-2$）
		$\sigma_1^2\neq\sigma_2^2$	独立	$n_1\neq n_2$	t′
				$n_1=n_2$	t（$df=n-1$）
		不考虑σ_1^2、σ_2^2是否相等	相关		t（$df=n-1$）
非正态分布	已知		独立、n_1与n_2都大于30（或50）		Z′
			相关、$n>30$（或50）		Z′
	未知		独立、n_1与n_2都大于30（或50）		Z′
			相关、$n>30$（或50）		Z′

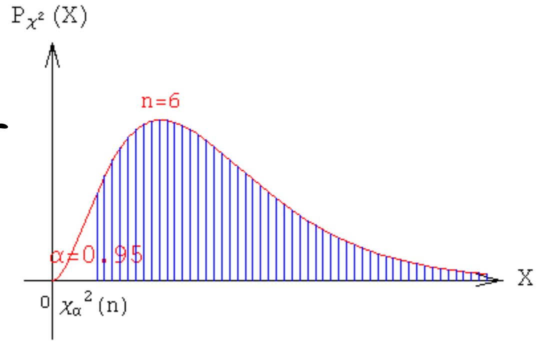
第四节 方差的差异检验

一、单总体方差的显著性检验

单总体方差的差异检验也称为样本方差与总体方差的差异显著性检验。由抽样分布的讨论知，从正态分布的总体中随机抽取容量为n的样本，其**样本方差与总体方差比值的分布为自由度为n-1 的χ^2 分布**。其检验公式为：

$$\chi^2 = \frac{nS^2}{\sigma^2}$$

$$\chi^2/n = \frac{S^2}{\sigma^2}$$ 小概率事件



当$\chi^2 > \chi^2_{\alpha/2}$或$\chi^2 < \chi^2_{(1-\alpha/2)}$时，$S^2$与$\sigma^2$差异显著，

当$\chi^2_{(1-\alpha/2)} < \chi^2 < \chi^2_{\alpha/2}$时，$S^2$与$\sigma^2$差异不显著。

一、单总体方差的显著性检验

例14 已知某省化学高考成绩服从正态分布，标准差为15分，从某校抽取35名学生，测其化学高考成绩标准差为13分，试问该校化学高考成绩与全省方差有无显著性差异。

解：①建立假设: $H_0:\sigma_1^2=\sigma_0^2, H_1:\sigma_1^2\neq\sigma_0^2$

②计算统计量: $\chi^2=\dfrac{nS^2}{\sigma^2}=\dfrac{35\times13^2}{15^2}=\dfrac{5915}{225}=26.29$

③判断结果:查χ^2值表, 当$\alpha=0.05$, df=35-1=34, 得 $\chi^2_{0.05/2}=\chi^2_{0.025}=53$ (用内插法近似求)，$\chi^2_{1-0.05/2}=20.6$，因为$20.6<26.69<53$, 所以$p>0.05$, 差异不显著

二、双总体方差的差异显著性检验

两样本方差的差异显著性检验的目的是想了解样本抽自的两个总体的方差的一致情况，或两个总体方差是否相等，因此也称为方差的齐性检验。

㈠独立样本

设有两个正态总体X和Y, $X \sim N(\mu_1, \sigma_1^2)$, $Y \sim N(\mu_2, \sigma_2^2)$, 且X和Y相互独立,从X和Y两总体中抽取样本容量分别为n_1,n_2的两个样本,计算其方差分别为S_1^2 和S_2^2, 假如$\sigma_1^2 = \sigma_2^2$ 那么$\sigma_1^2 / \sigma_2^2 = 1$。当两样本的总体方差未知，则可用总体方差的无偏估计S_{n1-1}^2 和S_{n2-1}^2来代替σ_1^2 和 σ_2^2 。这两个估计量的比的抽样分布服从$df_1=n_1-1$，$df_2=n_2-1$的 F分布。

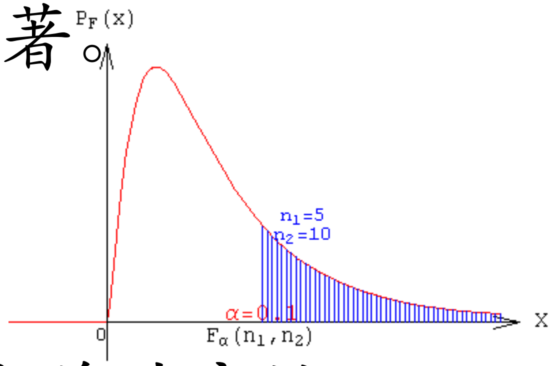
二、双总体方差的差异显著性检验

即（假如$\sigma_1^2=\sigma_2^2$ 根据公式6-14推导出6-15）

$$F=\frac{S_{n_1-1}^2/\sigma_1^2}{S_{n_2-1}^2/\sigma_2^2}=\frac{S_{n_1-1}^2}{S_{n_2-1}^2}\quad(df_1=n_1-1, df_2=n_2-1)$$

当$F>F_{\alpha/2}$或$F<F_{(1-\alpha/2)}$时，σ_1^2和σ_2^2差异显著，

当$F_{(1-\alpha/2)}<F<F_{\alpha/2}$时，$\sigma_1^2$和$\sigma_2^2$差异不显著



$F_{(1-\alpha/2)}$与$F_{\alpha/2}$互为倒数，即$F_{\alpha/2}=1/F_{(1-\alpha/2)}$

为了方便查F值表而不用计算倒数，在方差的齐性检验时，其检验公式为：

$$F=\frac{S_{大}^2}{S_{小}^2}$$

例题

例15 某区幼儿园对幼儿进行健康检查，从中随机抽取20名4 岁幼儿，其中男生12名，女生8名，经计算得男生体重标准差为1.75公斤,女生体重的标准差为1.73公斤,试问该区4岁男女幼儿体重的方差差异是否显著。

例15计算

解：①建立假设:$H_0:\sigma_1^2=\sigma_2^2, H_1:\sigma_1^2\neq\sigma_2^2$

②计算统计量:

$$S_{n_1-1}^2=\frac{n_1}{n_1-1}S_1^2=\frac{12}{11}\times1.75^2=3.34$$

$$S_{n_2-1}^2=\frac{n_2}{n_2-1}S_2^2=\frac{8}{7}\times1.73^2=3.42$$

③判断结果:

$$F=\frac{S_{n_2-1}^2}{S_{n_1-1}^2}=\frac{3.42}{3.34}=1.02$$

查F值得:$F_{(7,10).05/2}=3.95, F_{(7,12).05/2}=3.61$

由内插法得:$F_{(7,11).05/2}=3.78, \because 1.02<3.78,$

$\therefore P>0.05$,差异不显著.

例题

例16 本章第一节例10进行方差的差异显著性检验。

解：①建立假设：$H_0:\sigma_1^2=\sigma_2^2, H_1:\sigma_1^2\neq\sigma_2^2$

②计算统计量:

$$S_{n_1-1}^2=\frac{n_1}{n_1-1}S_1^2=\frac{22}{21}\times 8^2=67.05$$

$$S_{n_2-1}^2=\frac{n_2}{n_2-1}S_2^2=\frac{18}{17}\times 7.6^2=61.16$$

$$F=\frac{S_{n_1-1}^2}{S_{n_2-1}^2}=\frac{67.05}{61.16}=1.10$$

③判断结果:

查F值表得$:F_{(20,15)0.05/2}=2.76, F_{(24,15)0.05/2}=2.70$

$F_{(20,20)0.05/2}=2.46, F_{(24,20)0.05/2}=2.41,$ 由内插法得:

$F_{(21,17)0.05/2}=2.58, \because 1.1<2.58, \therefore P>0.05,$差异不显著.

二、双总体方差的差异显著性检验

(二) 相关样本

当两个样本相关时，两样本方差的比值就不再服从F分布，不能用F检验对方差的差异情况进行检验。但统计量：

$$t=\frac{S_1^2-S_2^2}{\sqrt{\frac{4S_1^2S_2^2\left(1-r^2\right)}{n-2}}}\quad\ldots(8-25)$$

服从自由度为n-2的t分布。因此可以对相关样本方差进行t检验。

例题

例19 在某次教学方法改革实验结束后，实验班与对照班（各班均为40人）的成绩的标准差分别为8.4分和12.6分，两班的相关系数为0.65,试问两班成绩的方差的差异是否显著。

解：①建立假设: $H_0:\sigma_1^2=\sigma_2^2;H_1:\sigma_1^2\neq\sigma_2^2$

②计算统计量:

$$t=\frac{S_1^2-S_2^2}{\sqrt{\frac{4S_1^2S_2^2\left(1-r^2\right)}{n-2}}}$$

$$=\frac{8.4^2-12.6^2}{\sqrt{\frac{4\times8.4^2\times12.6^2\times\left(1-0.65^2\right)}{40-2}}}=-3.38$$

例19(解续)

③判断结果:查t值表得:

$$t_{(38).01/2} = 2.704$$

$$\because \left|-3.38\right| > 2.704, \therefore p < 0.01$$

故两班成绩的方差差异极显著.

第五节 相关系数差异显著性检验

一、单总体相关系数差异显著性检验

㈠ 当ρ=o时

当ρ=o时，样本积差相关系数的抽样分布服从自由度为n-2的 **t** 分布，检验公式

$$t=\frac{r-\rho}{SE_r}=\frac{r-0}{SE_r}\ldots(8-26)$$

其中，标准误为：

$$SE_r=\frac{\sqrt{1-r^2}}{\sqrt{n-2}}$$

$$t=r\sqrt{\frac{n-2}{1-r^2}}$$

其他类型相关系数的显著性检验见p250-251

一、单总体相关系数差异显著性检验

例20 某校对各门学科进行统考，统考后从该校抽取42 名学生，计算他们语文与化学成绩的积差相关系数为0.34,试问语文与化学成绩是否真实存在相关。

解： 提出假设 H0： ρ=0， H1： ρ≠0

选择检验统计量 $t=\frac{r\sqrt{n-2}}{\sqrt{1-r^2}}=\frac{0.34\sqrt{40}}{\sqrt{1-0.34^2}}=2.2866$

统计决断： 根据df=42-2=40,查t值表得

$t_{(40)0.01}=2.704$, $t_{(40)0.05}=2.021$.

在0.01显著性水平拒绝零假设，而在0.05显著性水平下接受零假设

另一种方法：查积差相关系数临界表

- 根据df=40，查附表7，
- 从$\alpha=0.01$一列找到对应的积差相关系数临界值为0.393
- 从$\alpha=0.05$一列找到对应的积差相关系数临界值为0.304
- 而计算得到的$r=0.34$，大于表中$\alpha=0.05$的临界值，因此该接受该相关系数显著的结论。而它小于$\alpha=0.01$的临界值，因此该拒绝该相关系数显著的结论。
- **斯皮尔曼等级相关系数的检验就是查表做判断(附表9)**

一、单总体相关系数差异显著性检验

(二) ρ≠o时

当ρ≠o时，相关系数呈不同程度的偏态分布，只有样本足够大（n>500)时,相关系数的抽样分布才渐进正态分布。这一条件在实际中很难达到。这时，则需要用费舍法。检验的公式为

$$Z=\frac{Z_r-Z_\rho}{SE_r}\ldots(8-27)$$

Zr的标准误为:

$$SE_r=\frac{1}{\sqrt{n-3}}$$

$$Z=(Z_r-Z_\rho)\sqrt{n-3}$$

例题

例21有人估计,3~6岁儿童身高与体重的相关系数为0.80,从某幼儿园随机抽取3~6岁儿童40名,经测量,计算出他们的身高与体重的相关系数为0.68,试问在该幼儿园的测查结果能否推翻上述对身高与体重关系的估计。

解 H_0: $\rho=0$, H_1: $\rho\neq0$

0.8 的Zr为1.099

0.68 的Zr为0.829

$\alpha=0.05$, $Z_{\alpha/2}=1.96$

$$Z=(Z_r-Z_\rho)\sqrt{n-3}=(1.099-0.829)\sqrt{40-3}=1.6423<1.96$$

$\alpha=0.05$下，接受H_0，拒绝H_1

二、双总体相关系数差异显著性检验

㈠两样本相关系数分别由两组相互独立的被试得到

根据费舍Z函数进行转换，由于Z_r的分布近似正态，同样Z_{r_1}-Z_{r_2}的分布也是近似正态，检验的公式为：

$$Z=\frac{\left(Z_{r_1}-Z_{r_2}\right)-\left(Z_{\rho_1}-Z_{\rho_2}\right)}{SE_{D_r}}\ldots(8-32)$$

其中标准误为：

$$SE_{D_r}=\sqrt{\frac{1}{n_1-3}+\frac{1}{n_2-3}}$$

二、双总体相关系数差异显著性检验

例22 从某重点中学中抽取学生30名，从一般中学中抽取35名学生，分别计算他们的数学成绩与瑞文测验分数的相关系数为0.72和0.51,试问能否认为重点中学学生在数学成绩与瑞文推理能力之间比一般中学学生有更高的相关。

解 $H_0: r_1 \le r_2$, $H_1: r_1 > r_2$

0.72 的Zr为0.908

0.51 的Zr为0.563

$SE^2 = 1/27+1/32$， $SE = 0.2613$, $\alpha=0.05$, $Z_\alpha=1.64$

$$Z = \frac{0.908-0.563}{0.2613} = 2.4683 > 1.64$$

$\alpha=0.05$下，拒绝H_0，接受H_1

二、双总体相关系数差异显著性检验

㈡两样本相关系数由同一组被试求得

设r_{12}，r_{23}, r_{13}是由三列变量计算出的两两相关系数，若对r_{12}和r_{13}进行差异显著性检验时，样本相关系数的抽样分布服从自由度为n-3的t分布。其检验公式为：

$$t=\frac{(r_{12}-r_{13})\cdot\sqrt{(n-3)(1-r_{23})}}{\sqrt{2(1-r_{12}^2-r_{13}^2-r_{23}^2+2\cdot r_{12}\cdot r_{13}\cdot r_{23})}}\cdots(8-34)$$

二、双总体相关系数差异显著性检验

例23 某市统考结束后，随机抽取80名学生，分别计算出他们的数学成绩与物理成绩的相关系数为0.75,数学成绩与化学成绩的相关系数为0.54,物理成绩与化学成绩的相关为0.6,试问数学与物理和数学与化学的相关的差异是否显著。

例23计算

解：①建立假设：$H_0:\rho_{12}=\rho_{13}, H_1:\rho_{12}\neq\rho_{13}$

②计算统计量：

$$t=\frac{(r_{12}-r_{13})\cdot\sqrt{(n-3)(1-r_{23})}}{\sqrt{2(1-r_{12}-r_{13}-r_{23}+2\cdot r_{12}\cdot r_{13}\cdot r_{23})}}$$

$$=\frac{(0.75-0.54)\sqrt{(80-3)(1-0.6)}}{\sqrt{2(1-0.75^2-0.54^2-0.6^2+2\times0.75\times0.54\times0.6)}}$$

$$=\frac{1.165}{\sqrt{0.54}}=1.59$$

③判断结果:

查t值表得：$t_{(77).05/2}=1.99, \because 1.59<1.99,$

$\therefore P>0.05$，差异不显著.

第六节 比率的显著性检验

一、单总体比率的显著性检验

对样本比率与总体比率直接是否存在显著差异进行检验。

(1) 当$np \geq 5$或$nq \geq 5$时，样本比率$\hat{p}$的分布为渐近正态分布。检验的公式为

$$Z = \frac{\hat{p}-p_0}{SE_p}, \quad SE_p = \sqrt{\frac{p_0 q_0}{n}} \qquad \cdots(8-36)$$

确定检验形式：双侧，单侧

进行统计推断——查表寻找相应的临界值，比较Z与Z_α，从而确定该样本的p是否为小概率。

第六节 比率的显著性检验

例题24 已知某年某区高考升学率为75%，某校在这一年有300名学生参加了高考，最后有210人被高校录取，问该校的升学率与全区的升学率是否相同?

例24 计算

- 解：np>5, 提出假设 Ho: P′=0.75, H1: p′≠0.75
- 计算

$$Z = \frac{p' - p_0}{\sqrt{\frac{p_0 q_0}{n}}} = \frac{210/300 - 0.75}{\sqrt{\frac{0.75 \times 0.25}{300}}} = 2.00$$

- 统计决断：因为$Z=2>1.96=Z_{0.05/2}$, $p<0.05$，所以拒绝零假设，接受备择假设，即该校这一年的高考升学率与全区的升学率有显著差异，由实际数据来看，该校的这一年的升学率低于全区。

第六节 比率的显著性检验

一、单总体比率的显著性检验

(1) 当$np \le 5$ $(p<q)$时，可直接查附表13(1,2)，若$\hat{p}$落入p_0置信区间之外，可认为差异显著，否则不显著。

例题25 已知某区学习障碍儿童的比率为8%，通过调查得知某班45名学生中有学习障碍的学生有3人，问该班的学习障碍学生的比率与全区是否有差异?

例25计算

- np<5
- 通过查表得，比率为0.08所属总体比率0.95置信区间为2%~19%，包含了6.67%，所以要保留零假设，拒绝备择假设，也就是说该班学习障碍学生的比率与全区没有显著差异。

二、双总体比率差异显著性检验

- 总体比率差异的显著性检验是根据两个样本的比率来检验两个相应总体的比率是否存在显著性差异。由于样本性质不同，其检验方法也不同。

独立样本比率差异的显著性检验

当$n_1 p_1 \geq 5$，$n_2 p_2 \geq 5$时，统计量$\hat{p}_1 - \hat{p}_2 = D_p$的分布为正态分布。

1.若统计假设为： $H_0: p_1 - p_2 = p_D$， $H_1: p_1 - p_2 \neq p_D$

临界比率CR为：

$$Z = \frac{(\hat{p}_1 - \hat{p}_2) - p_D}{\sigma_{\hat{p}_1 - \hat{p}_2}} \qquad \sigma_{p_1 - p_2} = \sqrt{\frac{p_1 q_1}{n_1} + \frac{p_2 q_2}{n_2}}$$

2.若统计假设为： $H_0: p_1 = p_2 = p$， $H_1: p_1 \neq p_2$

临界比率CR为：

$$Z = \frac{(\hat{p}_1 - \hat{p}_2) - 0}{\sigma_{\hat{p}_1 - \hat{p}_2}} = \frac{\hat{p}_1 - \hat{p}_2}{\sigma_{\hat{p}_1 - \hat{p}_2}} \qquad \sigma_{p_1 - p_2} = \sqrt{\frac{(n_1 \hat{q}_1 + n_2 \hat{q}_2)(n_1 \hat{p}_1 + n_2 \hat{p}_2)}{n_1 n_2 (n_1 + n_2)}}$$

通过比较单、双侧，来判断比率差异是否显著。

例题26

- 例26 研究者在两所学校随机各抽取一个班学生进行一项视力检查，结果发现，甲校46名学生中患近视的有20人，乙校48名学生中患近视的有33人，问两校学生患近视的比率是否有显著性差异?

例题26 计算

- 解：提出假设　Ho: $p_1'=p_2'$　　H1 : $p_1' \neq p_2'$

计算检验统计量的值：n1 = 46, n2 = 48

p1=20/46=0.4348，q1 = 1-p1 = 0.5652

p2=33/48=0.6875，q2 = 1-p2 = 0.3125

$$Z=\frac{\hat{p}_1-\hat{p}_2}{\sqrt{\frac{(n_1\hat{q}_1+n_2\hat{q}_2)(n_1\hat{p}_1+n_2\hat{p}_2)}{n_1n_2(n_1+n_2)}}}=\frac{0.4348\text{-}0.6875}{\sqrt{\frac{(20+33)(26+15)}{46\times48(46+48)}}}=2.47$$

统计决断：因为Z=2.47>1.96=$Z_{0.05/2}$，p<0.05，所以拒绝零假设，接受备择假设，即两校学生患近视的比

例题27

- 例27 某厂质量检验人员认为该厂一车间的产品一级品的比率比二车间产品一级品的比率大5%，现从一车间和二车间分别抽出2个独立随机样本，得到如下数据: $n_1=150$, 其中一级品为113；$n_2 = 160$, 其中一级品为104.试根据这些数据检验质量研究人员的观点。（设$\alpha=0.05$）

例27计算

- 建立假设 Ho: $p1-p2 \le 0.05$ H1: $p1-p2 > 0.05$
- $p1=113/150=0.753$, $p2 = 104/160=0.650$

$$Z=\frac{(\hat{p}_1-\hat{p}_2)-0.05}{\sqrt{\frac{\hat{p}_1\hat{q}_1}{n_1}+\frac{\hat{p}_2\hat{q}_2}{n_2}}}=1.034$$

- 这个是右侧检验，$\alpha=0.05$, $Z_{0.05}=1.645>1.034$，故不能否定原假设，也就是说，不支持该厂质量检测人员的观点是正确的。

独立样本比率差异的显著性检验

当 $n_1p_1 \prec 5,\ n_2p_2 \prec 5$
差异显著性检验应用精确概率检验法。具体见第十章

相关样本比率差异的显著性检验

- 当两样本相关时，同一组被试在前后两次实验中，观察的两个项目又相同，这样便可得到前后两个项目的结果，据这两个结果所计算出来的两样本比率，就称作相关样本比率。

例题28

- 例28 某校120名学生期末代数测验之后，让他们在寒假里独立完成教师编选的代数练习题，开学初进行同类题目的测验，两次测验结果见下面频数表，问学生独立完成教师编选的代数练习题，对提高代数成绩是否有显著效果?

		第二次测验		合计
		良	非良	
第一次测验	良	48 (a)	14 (b)	62 (a+b)
	非良	22 (c)	36 (d)	58 (c+d)
合计		70 (a+c)	50 (b+d)	120 (n)

例28 计算

- 根据表中的符号，可以把两次测验成绩良好的比率之差表示为
- $p_1-p_2 = (a+b)/n - (a+c)/n = (b-c)/n$
- 也就是两次成绩良好比率之差只涉及两次不同成绩b和c的差异是否显著。
- 设两次测验中成绩不一致的数量为K=b+c，实验结果b、c可视为二项分布（b可称为成功的次数，c视为失败的次数）。

例28 计算

- 若b=c，则p=q=0.5(即两次实验结果的比率差异各为0.5，无差异)，若Kp≥5，按二项分布可知（近似正态分布）：

  $\mu=Kp=(b+c)/2,\ \sigma^2=Kpq=(b+c)/4$

- 如果每次取样，b和c落入b(b,K,0.5)的置信区间内，则可推论说两次实验结果的比率差异不显著；否则，则说明不是来自于p=q=0.5的总体，而是取自p≠q≠0.5的总体，即比率差异显著。

例28 计算

- 根据显著性检验公式

$$Z=\frac{p-p_0}{\sqrt{\frac{p_0q_0}{K}}}=\frac{K\left(p-p_0\right)}{\sqrt{Kp_0q_0}}=\frac{K\left(p-0.5\right)}{\sqrt{Kp_0q_0}}$$

$$=\frac{b-\frac{1}{2}(b+c)}{\sqrt{\left(b+c\right)\frac{1}{2}\times\frac{1}{2}}}=\frac{b-c}{\sqrt{b+c}}=\frac{14-22}{\sqrt{14+22}}=-1.33$$

- 因为|Z|=1.33<1.96，p>0.05，所以保留零假设，即寒假里独立完成教师编选的代数练习题对提高代数成绩无显著效果。

例29

- 例29 有研究者通过在一年前后对同一群学生各进行了一次调查来了解某校图书馆服务质量的改善情况。两次调查的结果如下，问两次调查所获得的满意率是否有显著差异?

		第二次调查		合计
		满意	不满意	
第一次调查	满意	18	10	28
	不满意	32	12	44
合计		50	22	72

例题29 计算

- 根据表中的数据得知： b=10, c=32
- 将它们代入公式中，

$$Z=\frac{b-c}{\sqrt{b+c}}=\frac{10-32}{\sqrt{10+32}}=-3.40$$

- 因为|Z|=3.4>2.58, p<0.01, 所以拒绝零假设，得出结论：一年后学生们对该校图书馆服务的满意率有了极显著的提高。

练习题

观察次数		戒烟前	
		吸烟	不吸烟
戒烟后	吸烟	65	15
	不吸烟	25	55

当Kp<5, 或K<10时

- 不能用正态分布概率计算，需按二项分布计算p(x>b)（或p(x>c)）以上的概率和来得到临界比率，然后按照单双侧来判断差异是否显著。
- 举例： P259 例题

青山不改，绿水长流

江湖再见！