学习目标

1. 了解假设检验的基本思想
2. 掌握假设检验的步骤
3. 对实际问题作假设检验
4. 利用置信区间进行假设检验
5. 利用*P* - 值进行假设检验

假设问题的提出

什么是假设(hypothesis)

- 对总体参数的的数值所作的一种陈述
  - 总体参数包括**总体均值**、**比例**、**方差**等
  - 分析**之前**必需陈述

什么是假设检验?
(hypothesis testing)

1.事先对总体参数或分布形式作出某种假设，然后利用样本信息来判断原假设是否成立

2.有参数假设检验和非参数假设检验

3.采用逻辑上的反证法，依据统计上的小概率原理

提出原假设和备择假设

- ➡ **什么是原假设？(null hypothesis)**

1. 待检验的假设，又称“0假设”
2. 研究者想收集证据予以反对的假设
3. 总是有等号 =, ≤ 或 ≥
4. 表示为 H_0
   - H_0：μ = 某一数值
   - 指定为 = 号，即 ≤ 或 ≥
   - 例如, H_0：$\mu = 3190$（克）

提出原假设和备择假设

➡ 什么是备择假设？(alternative hypothesis)

1. 与原假设对立的假设，也称“研究假设”
2. 研究者想收集证据予以支持的假设总是有不等号: $\neq$, $<$ 或 $>$
3. 表示为 H_1
   - H_1 : μ <某一数值，或μ >某一数值
   - 例如, H_1 : μ < 3910(克)，或μ >3910(克)

假设检验中的两类错误
（决策风险）

假设检验中的两类错误

- **1.第一类错误（弃真错误）**
  - 原假设为真时拒绝原假设
  - 会产生一系列后果
  - 第一类错误的概率为α
    - 被称为显著性水平
- **2.第二类错误（取伪错误）**
  - 原假设为假时接受原假设
  - 第二类错误的概率为β (Beta)

假设检验的流程

- 提出假设
- 确定适当的检验统计量
- 规定显著性水平α
- 计算检验统计量的值
- 作出统计决策

确定适当的检验统计量

➡ 什么是检验统计量？

1. 用于假设检验决策的统计量
2. 选择统计量的方法与参数估计相同，需考虑
   - 是大样本还是小样本
   - 总体方差已知还是未知
3. 检验统计量的基本形式为

$$Z = \frac{\bar{X} - \mu_0}{\sigma / \sqrt{n}}$$

规定显著性水平α
(significant level)

➡ 什么是显著性水平？

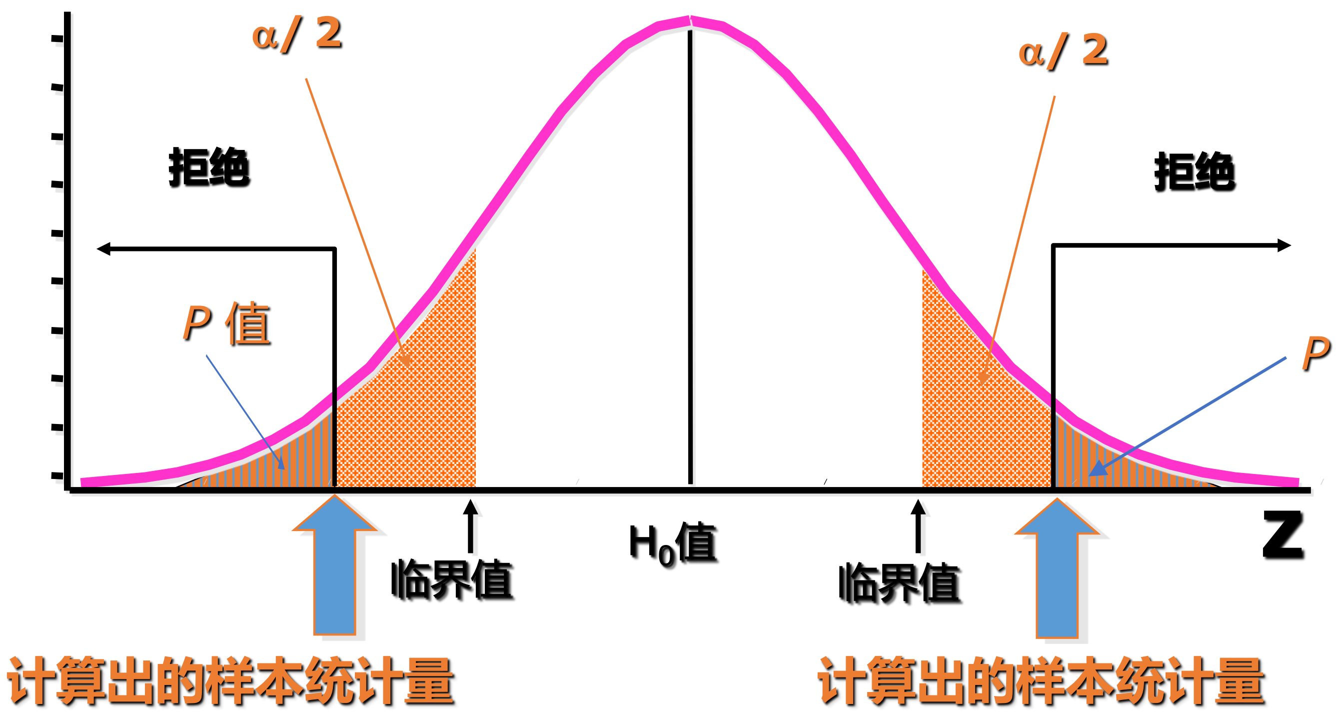
- 1. 是一个概率值
- 2. 原假设为真时，拒绝原假设的概率
  - 被称为抽样分布的拒绝域
- 3. 表示为 α (alpha)
  - 常用的 α 值有0.01, 0.05, 0.10
- 4. 由研究者事先确定

作出统计决策

1. 计算检验的统计量
2. 根据给定的显著性水平α，查表得出相应的临界值z_α或$z_{\alpha/2}$， t_α或$t_{\alpha/2}$
3. 将检验统计量的值与α水平的临界值进行比较
4. 得出拒绝或不拒绝原假设的结论

举个栗子

例：某食品厂要求生产食品的装袋重量为100克。作为质检部门，现需要检验新生产出的一批食品装袋重量是否合格。于是抽取其中50袋作为样本，计算得到样本均值为80克，样本均值的标准差为5克。问若取95%置信水平，这批食品装袋重量是否合格？

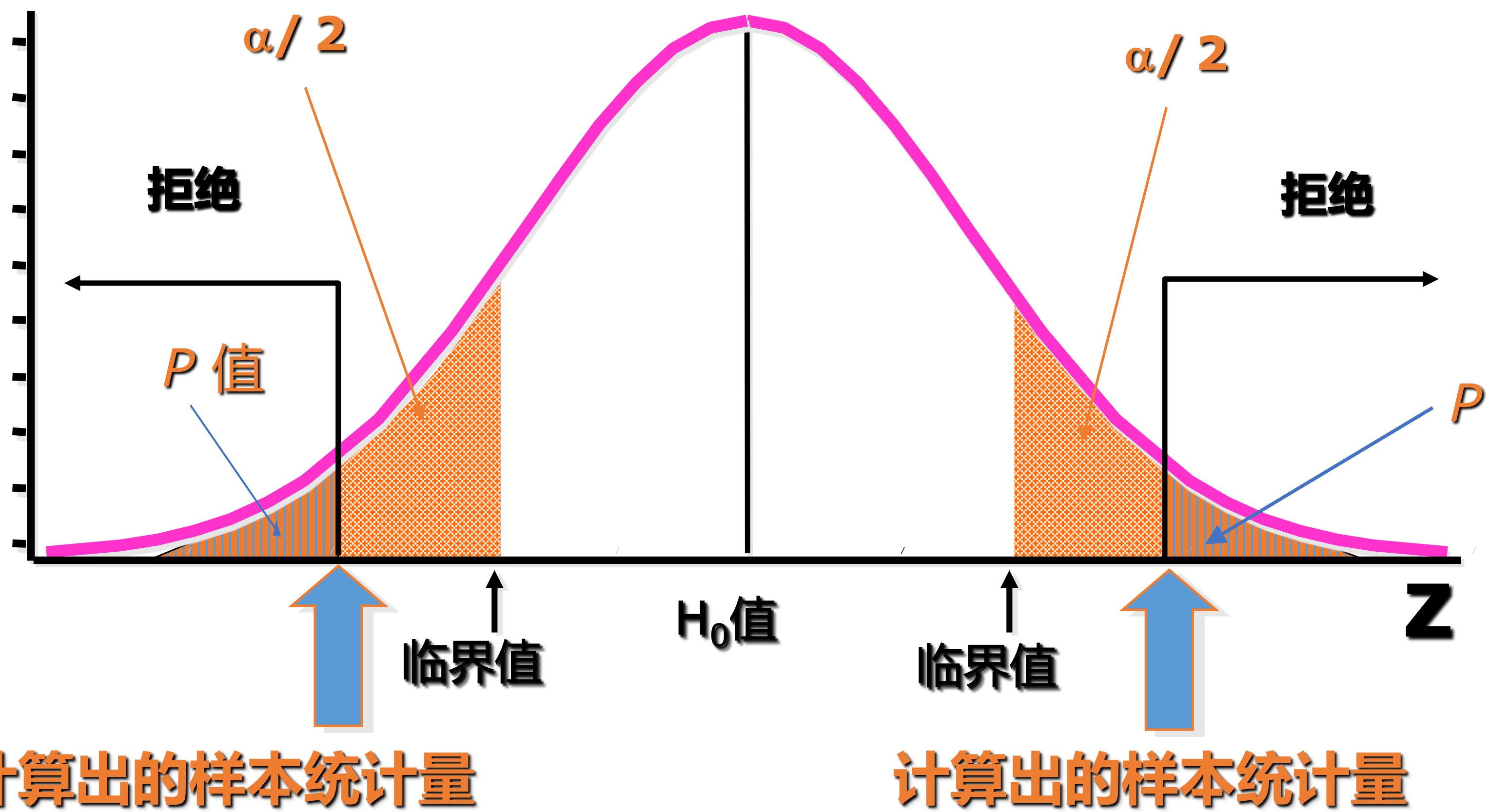


利用P值进行决策

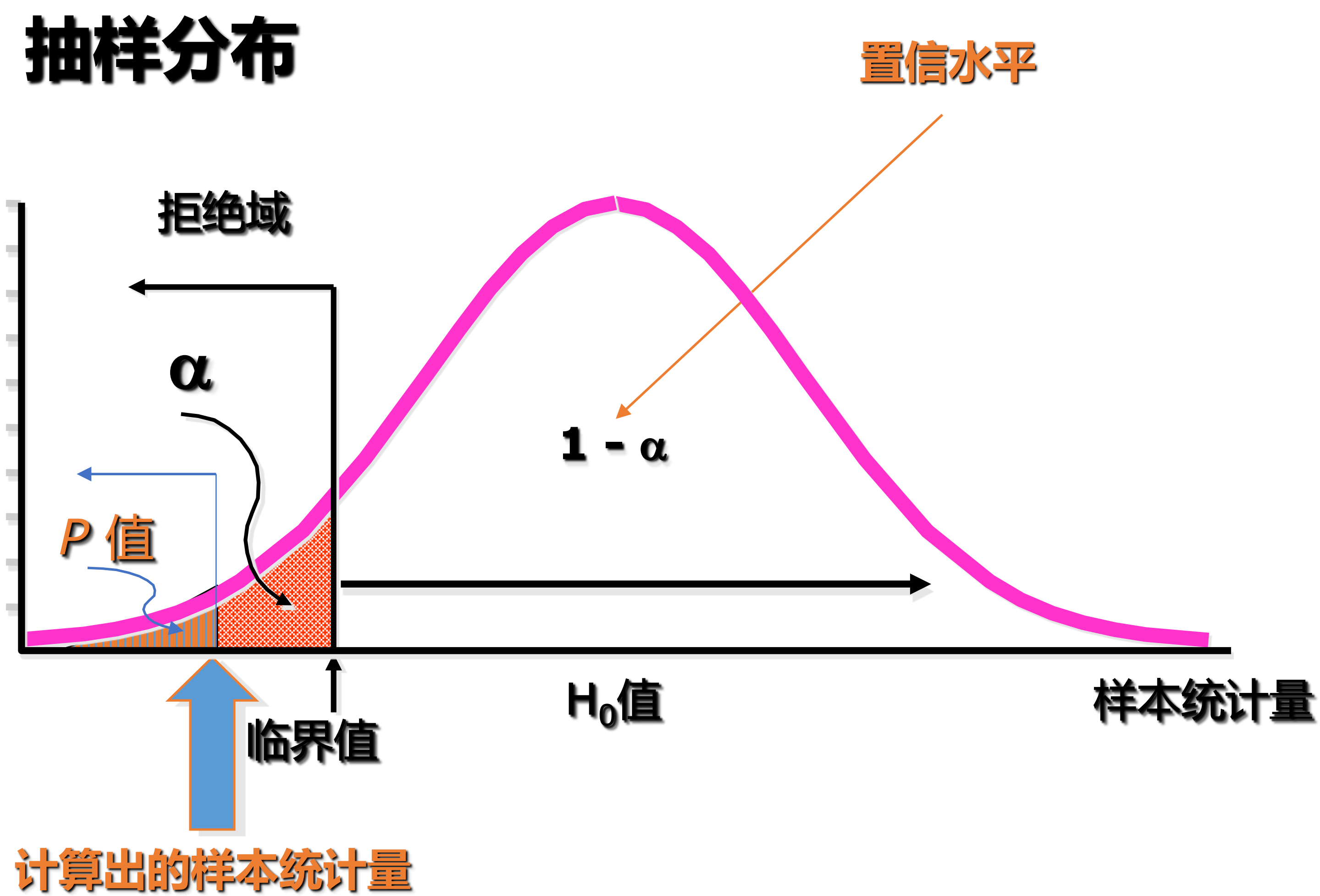
什么是*P* 值?
(*P*-value)

1. 是一个概率值
2. 如果原假设为真，*P*-值是抽样分布中大于或小于样本统计量的概率
   - 左侧检验时，*P*-值为曲线上方***小于等于***检验统计量部分的面积
   - 右侧检验时，*P*-值为曲线上方***大于等于***检验统计量部分的面积
3. 被称为观察到的(或实测的)显著性水平
   - H_0 能被拒绝的最小值

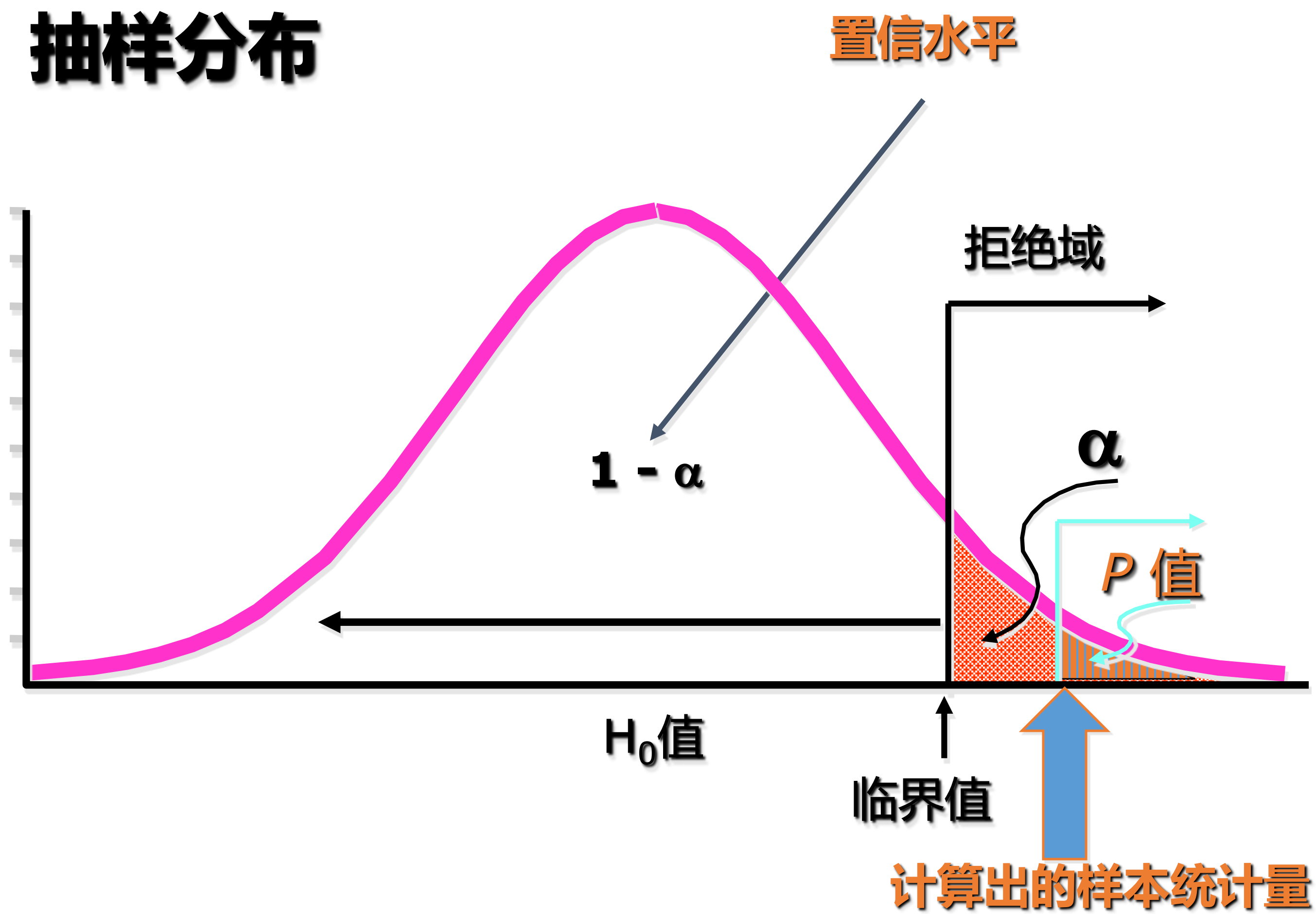
双侧检验的*P* 值



左侧检验的*P* 值



右侧检验的P 值



利用 *P* 值进行检验
(决策准则)

1. 单侧检验
   - 若p-值 > α, 不拒绝 H_0
   - 若p-值 < α, 拒绝 H_0
2. 双侧检验
   - 若p-值 > $\alpha/2$, 不拒绝 H_0
   - 若p-值 < $\alpha/2$, 拒绝 H_0

一个总体参数的检验

- 一个总体
  - 均值
    - Z 检验（单尾和双尾）
    - t 检验（单尾和双尾）
  - 比例
    - Z 检验（单尾和双尾）
  - 方差
    - χ^2检验（单尾和双尾）

总体均值检验

总体均值的检验
(检验统计量)

总体σ 是否已知？

是

z 检验

$$Z = \frac{\overline{X} - \mu_0}{\sigma / \sqrt{n}}$$

否

样本量n

大

z 检验

$$Z = \frac{\overline{X} - \mu_0}{S / \sqrt{n}}$$

小

用样本标准差S代替

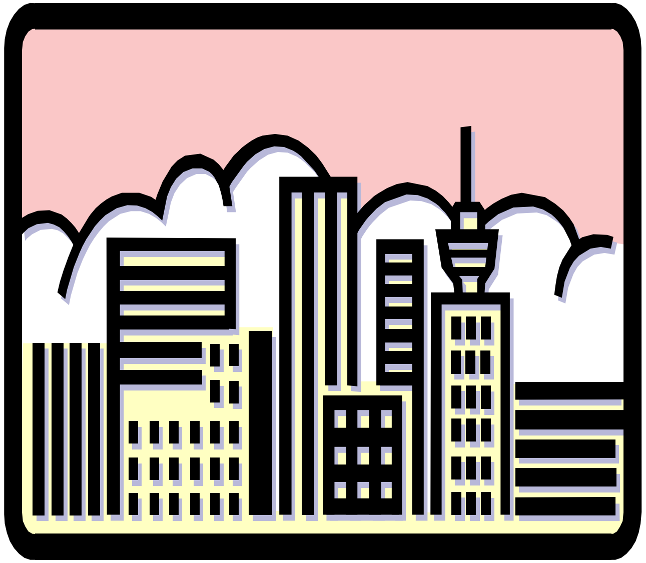
t 检验

$$t = \frac{\overline{X} - \mu_0}{S / \sqrt{n}}$$

σ^2 未知小样本均值的检验 (例题分析)

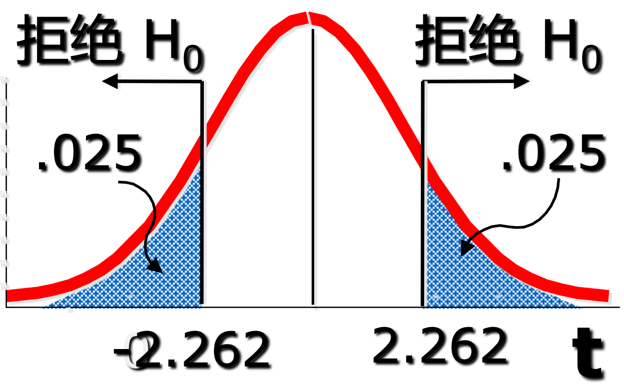
- 【例】某机器制造出的肥皂厚度为5cm，今欲了解机器性能是否良好，随机抽取10块肥皂为样本，测得平均厚度为5.3cm，标准差为0.3cm，试以0.05的显著性水平检验机器性能良好的假设。

双侧检验



σ^2 未知小样本均值的检验
(例题分析)

- H_0: $\mu = 5$
- H_1: $\mu \neq 5$
- $\alpha = 0.05$
- $df = 10 - 1 = 9$
- 临界值(s):



检验统计量:

$$t = \frac{\bar{x} - \mu_0}{s/\sqrt{n}} = \frac{5.3 - 5}{0.3/\sqrt{10}} = 3.16$$

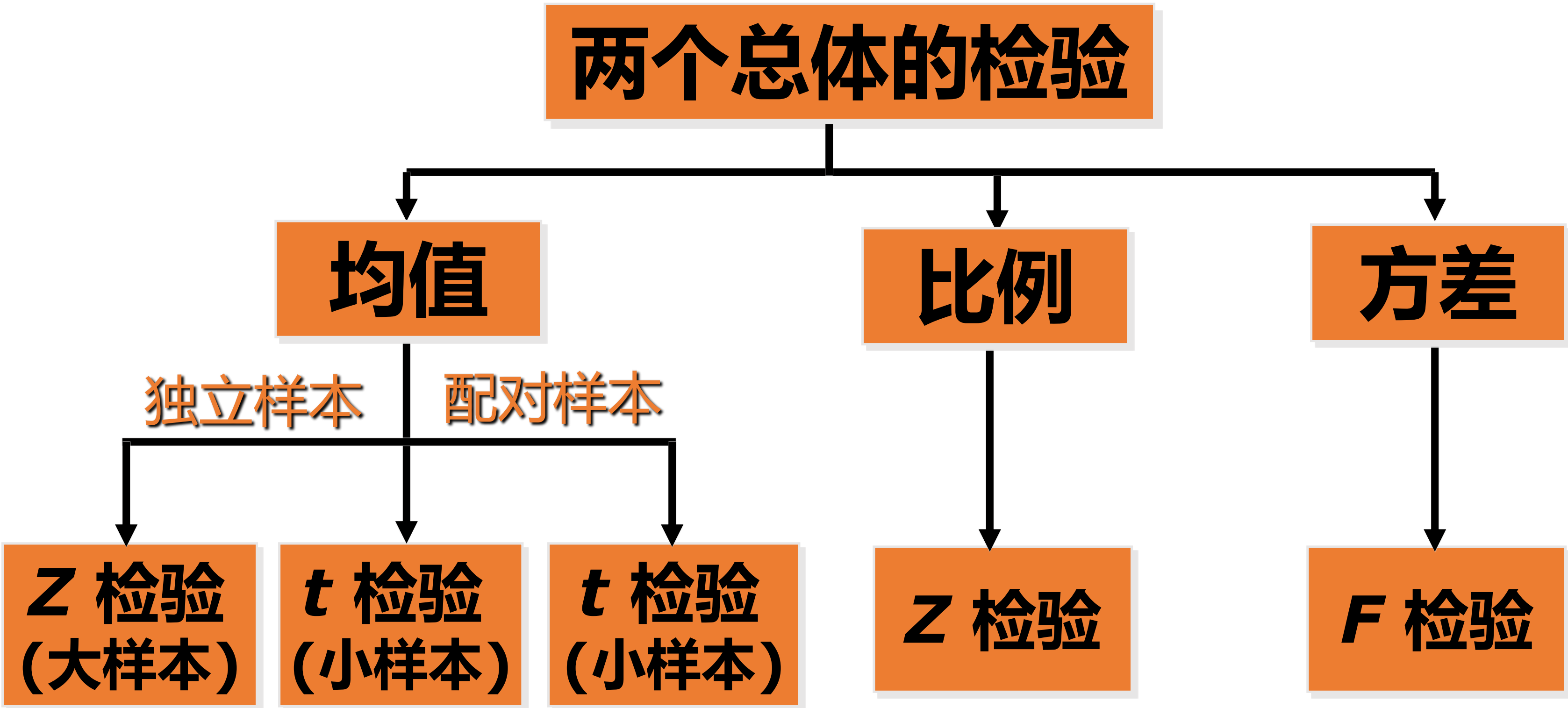
决策:

在 $\alpha = 0.05$的水平上拒绝H_0

结论:

说明该机器的性能不好

两个正态总体参数的检验



匹配样本的 t 检验

(例题分析)

- 【例】一个以减肥为主要目标的健美俱乐部声称，参加其训练班至少可以使减肥者平均体重减重**8.5**kg以上。为了验证该宣称是否可信，调查人员随机抽取了**10**名参加者，得到他们的体重记录如下表：

训练前	94.5	101	110	103.5	97	88.5	96.5	101	104	116.5
训练后	85	89.5	101.5	96	86	80.5	87	93.5	93	102

在 $\alpha = 0.05$ 的显著性水平下，调查结果是否支持该俱乐部的声称？

单侧检验

配对样本的 t 检验
(例题分析)

样本差值计算表		
训练前	训练后	差值D_i
94.5	85	9.5
101	89.5	11.5
110	101.5	8.5
103.5	96	7.5
97	86	11
88.5	80.5	8
96.5	87	9.5
101	93.5	7.5
104	93	11
116.5	102	14.5
合计	—	98.5

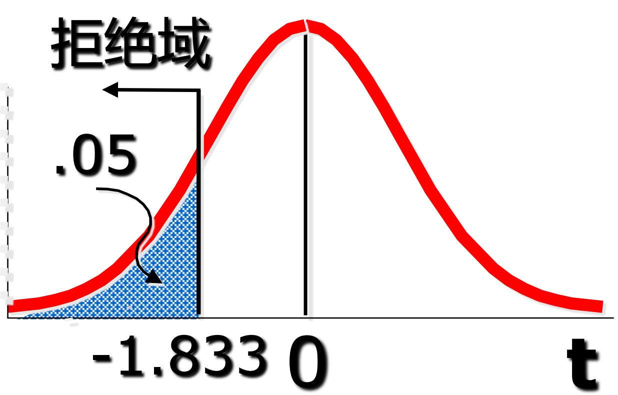
配对样本的 t 检验
(例题分析)

差值均值 $$\bar{x}_D = \frac{\sum_{i=1}^{n} D_i}{n_D} = \frac{98.5}{10} = 9.85$$

差值标准差 $$s_D = \sqrt{\frac{\sum_{i=1}^{n}(D_i - \bar{x}_D)^2}{n_D - 1}} = \sqrt{\frac{43.525}{10-1}} = 2.199$$

配对样本的 t 检验
(例题分析)

- H_0: $\mu_1 - \mu_2 \geq 8.5$
- H_1: $\mu_1 - \mu_2 < 8.5$
- $\alpha = 0.05$
- df = 10 - 1 = 9
- 临界值(s):



检验统计量:

$$t = \frac{\bar{x}_D - D_0}{s_D / \sqrt{n_D}} = \frac{9.85 - 8.5}{2.199 / \sqrt{10}} = 1.9413$$

决策:

在 $\alpha = 0.05$的水平上不拒绝H_0

结论:

不能认为该俱乐部的宣称不可信