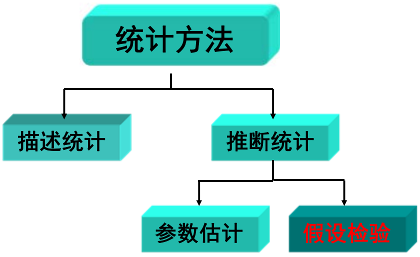
第五章 假设检验

第五章 假设检验

假设检验在统计方法中的地位



学习重点

1. 了解假设检验的基本思想
2. 掌握假设检验的步骤
3. 能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

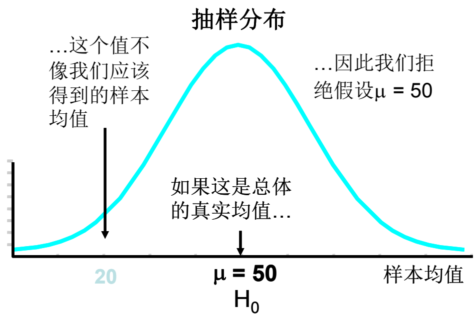
第一节 假设检验的一般问题

一. 假设检验的概念
二. 假设检验的步骤
三. 假设检验中的小概率原理
四. 假设检验中的两类错误
五. 双侧检验和单侧检验

什么是假设检验?

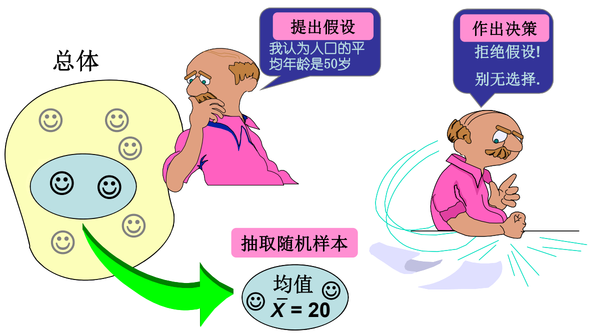
1. 概念
   - 事先对总体参数或分布形式作出某种假设
   - 然后利用样本信息来判断原假设是否成立
2. 类型
   - 参数假设检验
   - 非参数假设检验
3. 特点
   - 采用逻辑上的反证法
   - 依据统计上的小概率原理

假设检验的基本思想



假设检验的过程

（提出假设→抽取样本→作出决策）



假设检验的步骤

- 提出原假设和备择假设
- 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

例：某动物实验，要求动物平均体重$\mu_0 = 10.00g$，若μ < $10.00g$需再饲养，若μ > $10.00g$则应淘汰。动物体重服从正态分布$N(\mu,\sigma^2)$的随机变量。已知总体标准差$\sigma = 0.40g$，但总体平均数 μ 未知。为了得出对总体平均数 μ 的推断，从动物群体中，独立随机地抽取含量为 n 的样本，通过样本平均数推断总体平均数。

设 H_0: $\mu = \mu_0$

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.关于总体参数的假设

3.总是有等号： =, ≤ 或 ≥

4.表示为 H_0

例如, H_0：μ = 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

2.总是有不等号: ≠, < 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)

例：某动物实验，要求动物平均体重$\mu_0 = 10.00g$，若$\mu < 10.00g$需再饲养，若$\mu > 10.00g$则应淘汰。动物体重服从正态分布$N(\mu, \sigma^2)$的随机变量。已知总体标准差$\sigma = 0.40g$，但总体平均数 μ 未知。为了得出对总体平均数 μ 的推断，从动物群体中，独立随机地抽取含量为 n 的样本，通过样本平均数推断总体平均数。

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

确定适当的检验统计量

➡ 什么检验统计量？

1. 用于假设检验问题的统计量
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

例：某动物实验，要求动物平均体重$\mu_0 = 10.00g$，若$\mu < 10.00g$需再饲养，若$\mu > 10.00g$则应淘汰。动物体重服从正态分布$N(\mu, \sigma^2)$的随机变量。已知总体标准差$\sigma = 0.40g$，但总体平均数 μ 未知。为了得出对总体平均数 μ 的推断，从动物群体中，独立随机地抽取含量为 n 的样本，通过样本平均数推断总体平均数。

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

选用 Z 统计量　$z = \dfrac{\bar{x} - \mu_0}{\sigma/\sqrt{n}}$

假设检验中的小概率原理

➡ 什么小概率？

1.在一次试验中，一个几乎不可能发生的事件发生的概率。在一次试验中小概率事件一旦发生，我们就有理由拒绝原假设。

2. 小概率应该是接近0的一个数。

☆ 显著性水平α

规定显著性水平α

➡ 什么显著性水平？

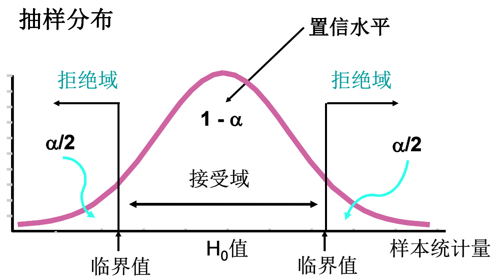
1.是一个概率值

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

双侧检验
（显著性水平与拒绝域 ）



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$Z_{\alpha/2}$
3. 将检验统计量的值与α水平的临界值进行比较
4. 得出接受或拒绝原假设的结论

☆ 显著性检验

例：某动物实验，要求动物平均体重$\mu_0 = 10.00g$，若$\mu < 10.00g$需再饲养，若$\mu > 10.00g$则应淘汰。动物体重服从正态分布$N(\mu, \sigma^2)$的随机变量。已知总体标准差$\sigma = 0.40g$，但总体平均数 μ 未知。从该动物群体中抽出含量n=10的样本，计算出样本平均数$\bar{x} = 10.23g$。

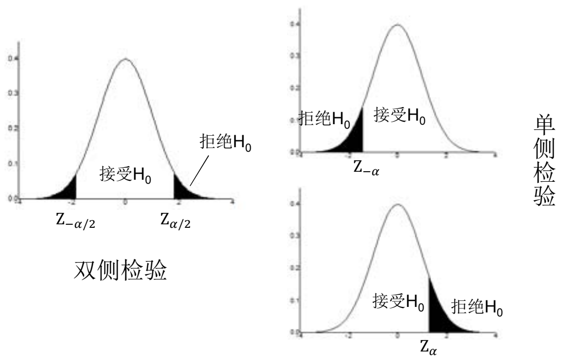
设 H_0: $\mu = \mu_0$ H_1: $\mu \neq \mu_0$

Z统计量 $Z = \dfrac{\bar{x} - \mu_0}{\sigma/\sqrt{n}} = 1.82$

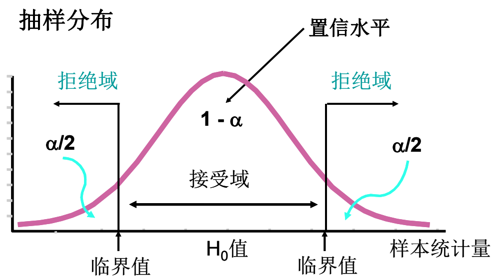
显著性水平$\alpha = 0.05$， $Z_{\alpha/2} = 1.96$

由于 -1.96<1.86<1.96，该检验不显著。

双侧检验与单侧检验



双侧检验
（显著性水平与拒绝域）



双侧检验
（例子）

该企业生产的零件平均长度是4厘米吗?

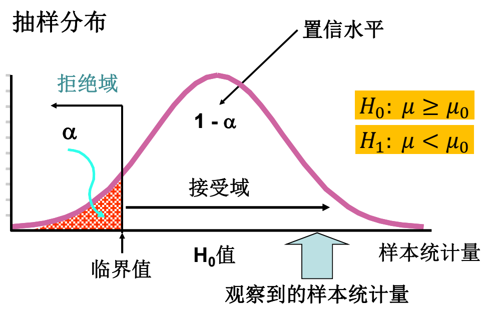
(属于决策中的假设)

- 提出原假设: H_0: $\mu = 4$
- 提出备择假设: H_1: $\mu \neq 4$

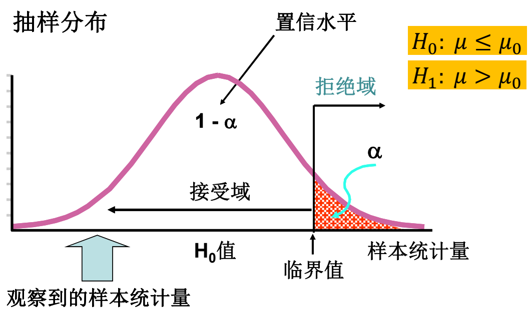
双侧检验与单侧检验
(假设的形式)

假设	研究的问题		
	双侧检验	左侧检验	右侧检验
H_0	$\mu = \mu_0$	$\mu \geq \mu_0$	$\mu \leq \mu_0$
H_1	$\mu \neq \mu_0$	$\mu < \mu_0$	$\mu > \mu_0$

左侧检验
（显著性水平与拒绝域）



右侧检验
（显著性水平与拒绝域）



单侧检验
（原假设与备择假设的确定）

➡ 检验*研究中的假设*

1. 将所研究的假设作为备择假设H_1
2. 将认为研究结果是无效的说法或理论作为原假设H_0。或者说，把希望(想要)证明的假设作为备择假设
3. 先确立备择假设H_1

单侧检验
（原假设与备择假设的确定）

➡检验*某项声明的有效性*

1. 将所作出的说明(声明)作为原假设
2. 对该说明的质疑作为备择假设
3. 先确立原假设H_0
   – 除非我们有证据表明“声明”无效，否则就应认为该“声明”是有效的

例：某动物实验，要求动物平均体重$\mu_0 = 10.00g$。动物体重服从正态分布$N(\mu, \sigma^2)$的随机变量。已知总体标准差$\sigma = 0.40g$，但总体平均数μ未知。从该动物群体中抽出含量n=10的样本，计算出样本平均数$\bar{x} = 10.23g$。已知这批动物实际饲养时间比根据以往经验所需饲养时间长得多，因此μ不可能小于μ_0。

设 H_0: $\mu = \mu_0$　　H_1: $\mu > \mu_0$

Z统计量 $Z = \dfrac{\bar{x} - \mu_0}{\sigma/\sqrt{n}} = 1.82$

$\alpha = 0.05$，单侧检验临界值 $Z_\alpha = 1.645$

由于1.86>1.645，拒绝H_0，接受H_1。

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
- 第一类错误的概率为α，被称为显著性水平

$\alpha = P(\text{拒绝}H_0\ |H_0\text{是正确的})$

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)

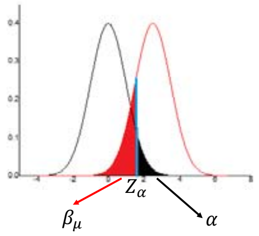
$\beta = P(\text{接受}H_0\ |H_0\text{是错误的})$

假设检验中的两类错误
（决策结果）

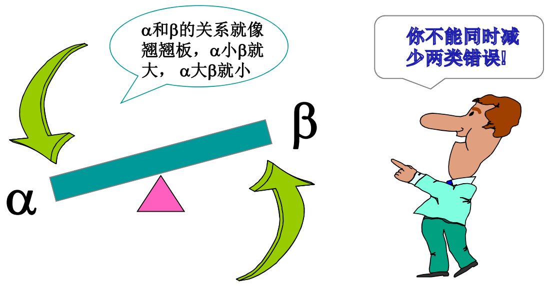
H_0 检验		
决策	实际情况	
	H_0为真	H_0为假
接受H_0	1 - α	第二类错误 (β)
拒绝H_0	第一类错误 (α)	功效 (1-β)

影响 β 错误的因素

1.总体参数的真值
真值越接近假设的参数值，犯II类错误概率越大
2.显著性水平 α
当 α 减少时增大
3.总体标准差 σ
当 σ 增大时增大
4.样本容量 n
当 n 减少时增大



α 错误和 β 错误的关系



第一节　假设检验的一般问题

一．假设检验的概念
二．假设检验的步骤
三．假设检验中的小概率原理
四．假设检验中的两类错误
五．双侧检验和单侧检验

第二节　一个正态总体的参数检验

一．总体方差已知时的均值检验
二．总体方差未知时的均值检验
三．总体比例的假设检验
四．总体方差的假设检验

检验的步骤

- 陈述原假设 H_0
- 陈述备择假设 H_1
- 选择显著性水平 α
- 选择检验统计量
- 给出临界值
- 搜集数据
- 计算检验统计量
- 进行统计推断
- 表述解释结果

总体方差已知时的均值检验

均值的 *Z* 检验
(σ^2 已知)

1. 假定条件

   总体服从正态分布

   若不服从正态分布，可用正态分布来近似($n \geq 30$)

2. 双侧检验：H_0: $\mu=\mu_0$；H_1: $\mu \neq \mu_0$

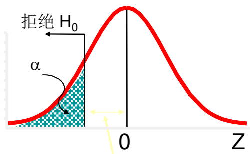
   单侧检验：备择假设有“<”或“>”符号

3. 使用*z*-统计量

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

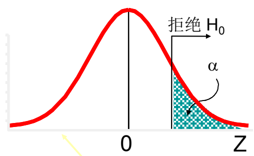
均值的单尾 Z 检验
（提出假设）

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



必须是*显著地*低于 μ_0，大的值满足H_0，不能拒绝

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



必须*显著地*大于μ_0，小的值满足 H_0，不能拒绝

均值的单尾Z检验

【例】已知豌豆籽粒重(mg)服从正态分布$N(377.2，3.3^2)$。在改善栽培条件后，随机抽取9粒，其籽粒重平均值$\bar{x}=379.2$。若标准差仍为3.3，问改善栽培条件是否显著提高了豌豆籽粒重？($\alpha=0.05$)

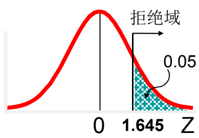
均值的单尾Z检验
（计算结果）

H_0: μ = 377.2

H_1: μ > 377.2

α = 0.05

临界值(s):



检验统计量:

$$Z=\frac{\bar{x}-\mu_0}{\sigma/\sqrt{n}}=\frac{379.2-377.2}{3.3/\sqrt{9}}$$

$= 1.82$

决策:

在 α = 0.05的水平上拒绝H_0

结论:

改善栽培条件显著提高了豌豆籽粒的粒重。

总体方差未知时的均值检验
(t 检验)

均值的 t 检验
(σ^2 未知)

1.假定条件

总体为正态分布

如果不是正态分布, 只有轻微偏斜和大样本 ($n \geq 30$)条件下

2.使用t 统计量

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

均值的双尾 t 检验

【例】已知玉米单交种"群单105"的平均穗重$\mu_0 = 300g$。喷药后，随机抽取9个果穗，其穗重分别为308g、305g、311g、298g、315g、300g、321g、294g、320g。问喷药后与喷药前的果穗重差异是否显著？

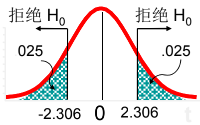
均值的双尾 t 检验

H0: μ = 300　H1: μ ≠ 300

$$\bar{x} = \frac{1}{9}\sum_{i=1}^{9} x_i = 308$$

$$s = \sqrt{\frac{\sum_{i=1}^{9} x_i^2 - \sum_{i=1}^{9} \bar{x}^2}{9-1}} = 9.62$$

临界值 $t_{0.025}(8)$:



检验统计量:

$$t = \frac{\bar{x} - \mu_0}{s/\sqrt{n}} = \frac{308 - 300}{9.62/\sqrt{9}}$$

= 2.49

决策:

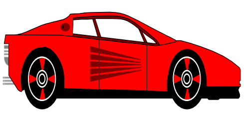
在 α = 0.05的水平上拒绝H_0

结论:

喷药前后果穗重的差异是显著的。

均值的单尾 t 检验
(实例)

属于检验声明有效性的假设!



【例】一个汽车轮胎制造商声称，某一等级的轮胎的平均寿命在一定的汽车重量和正常行驶条件下大于40000公里，对一个由20个轮胎组成的随机样本作了试验，测得平均值为41000公里，标准差为5000公里。已知轮胎寿命的公里数服从正态分布，我们能否根据这些数据作出结论，该制造商的产品同他所说的标准相符？(α = 0.05)

均值的单尾 t 检验
(计算结果)

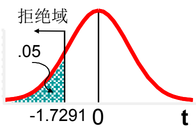
H0: μ ≥ 40000

H1: μ < 40000

α = 0.05

df = 20 - 1 = 19

临界值(s):



检验统计量:

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

$$= \frac{41000 - 40000}{5000/\sqrt{20}} = 0.894$$

决策:

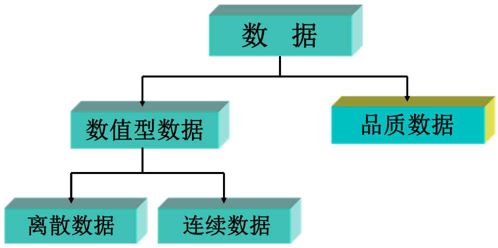
在α = 0.05的水平上接受H_0

结论:

有证据表明轮胎使用寿命显著地大于40000公里

总体比例的假设检验
（Z 检验）

适用的数据类型



一个总体比例的 Z 检验

1. 假定条件
   - 有两类结果
   - 总体服从二项分布
   - 可用正态分布来近似
2. 比例检验的 z 统计量

$$z=\frac{\hat{p}-p_0}{\sqrt{\frac{p_0(1-p_0)}{n}}}\sim N(0,1)$$

P_0为假设的总体比例

一个总体比例的 Z 检验
（实例）

【例】某研究者估计本市居民家庭的电脑拥有率为**30%**。现随机抽查了**200**的家庭，其中**68**个家庭拥有电脑。试问研究者的估计是否可信？(α = **0.05**)

一个样本比例的 Z 检验
（结果）

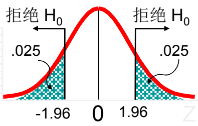
H_0: p = 0.3

H_1: $p \neq 0.3$

α = 0.05

n = 200

临界值(s):



检验统计量:

$$z=\frac{\hat{p}-p_0}{\sqrt{\frac{p_0(1-p_0)}{n}}}$$

$$=\frac{0.34-0.3}{\sqrt{\frac{0.3\times0.7}{200}}}=1.234$$

决策:

在α = 0.05的水平上接受H_0

结论:

有证据表明研究者的估计可信

总体方差的检验
(χ^2 检验)

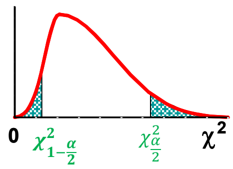
方差的卡方 (χ^2) 检验

1.检验一个总体的方差或标准差
2.假设总体服从正态分布
3.原假设为 H_0: $\sigma^2 = \sigma_0^2$
4.检验统计量

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

卡方(χ^2)分布的拒绝域

- 原假设 H_0: $\sigma = \sigma_0$
- 已知σ不可能小于σ_0，H_1: $\sigma > \sigma_0$
  拒绝域：$\chi^2 > \chi^2_\alpha$
- 已知σ不可能大于σ_0，H_1: $\sigma < \sigma_0$
  拒绝域：$\chi^2 < \chi^2_{1-\alpha}$
- H_1: $\sigma \neq \sigma_0$
  拒绝域：
  $\chi^2 < \chi^2_{1-\frac{\alpha}{2}}$ 和 $\chi^2 > \chi^2_{\frac{\alpha}{2}}$



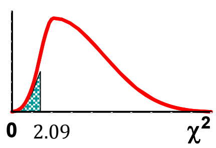
卡方 (χ^2)检验
实例

【例】某个混杂的小麦品种，株高标准差$\sigma_0 = 14cm$，经提纯后随机抽出10株，它们的株高分别为90cm、105cm、101cm、95cm、100cm、100cm、101cm、105cm、93cm、97cm，考虑提纯后的群体是否比原群体整齐？(α=**0.05**)

卡方 (χ^2) 检验

H0: σ = 14
H1: σ < 14
α = 0.01, *df*=10-1=9
临界值(s):

$\chi^2_{1-\alpha}(df) = 2.09$



统计量:

$$\chi^2 = \frac{(n-1)s^2}{\sigma^2} = \frac{\sum_{i=1}^{10}(x_i - \bar{x})^2}{\sigma^2} = 1.11$$

决策:

在 α = 0.01的水平上拒绝H_0

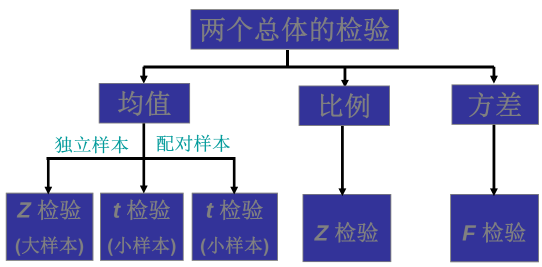
结论:

表明提纯后的株高比原群体高度整齐。

第三节 两个正态总体的参数检验

一. 两个总体方差齐性的检验
二. 两个总体均值之差的检验
三. 假设检验中相关样本的利用
四. 两个总体比例之差的检验

两个正态总体的参数检验



两个总体方差之比的F检验

1. 假定条件
   - 两个样本是独立的随机样本
   - 两个总体都是正态分布
2. 原假设：H_0: $\sigma_1 = \sigma_2$; 备择假设：H_1: $\sigma_1 \neq \sigma_2$
3. 检验统计量为

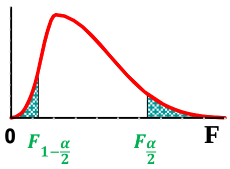
$$F = \frac{S_1^2}{S_2^2} \sim F(n_1 - 1, n_2 - 1)$$

F分布的拒绝域

- 原假设 H_0: $\sigma_1 = \sigma_2$
- H_1: $\sigma_1 > \sigma_2$ 拒绝域： $F > F_\alpha$
- H_1: $\sigma_1 < \sigma_2$ 拒绝域： $F < F_{1-\alpha}$
- H_1: $\sigma \neq \sigma_0$

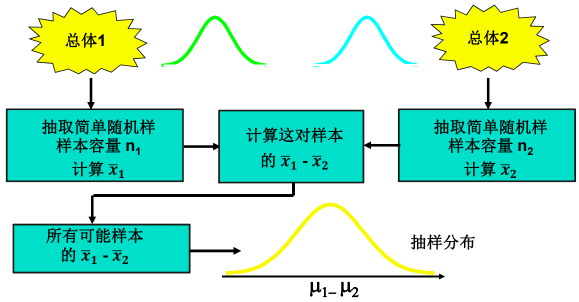
拒绝域：

$\chi^2 < \chi^2_{1-\frac{\alpha}{2}}$ 和 $\chi^2 > \chi^2_{\frac{\alpha}{2}}$



两个独立样本的均值检验

两个独立样本之差的抽样分布



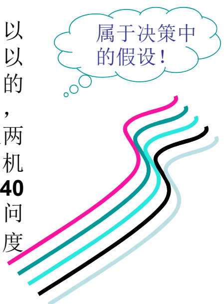
两个总体均值之差的Z检验 (σ_1^2、σ_2^2 已知)

1. 假定条件
   - 两个样本是独立的随机样本
   - 两个总体都是正态分布
   - 若不是正态分布，可以用正态分布来近似($n_1 \geq 30$ 和 $n_2 \geq 30$)
2. 原假设：H_0: $\mu_1 - \mu_2 = 0$; 备择假设：H_1: $\mu_1 - \mu_2 \neq 0$
3. 检验统计量为

$$z = \frac{(\bar{x}_1 - \bar{x}_2) - (\mu_1 - \mu_2)}{\sqrt{\frac{\sigma_1^2}{n_1} + \frac{\sigma_2^2}{n_2}}} \sim N(0,1)$$

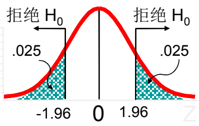
两个总体均值之差的Z检验

【例】有两种方法可用于制造某种以抗拉强度为重要特征的产品。根据以往的资料得知，第一种方法生产出的产品其抗拉强度的标准差为**8**公斤，第二种方法的标准差为**10**公斤。从两种方法生产的产品中各抽取一个随机样本，样本容量分别为n_1=**32**，n_2=**40**，测得 x_2= **50**公斤， x_1= **40**公斤。问这两种方法生产的产品平均抗拉强度是否有显著差别？ (α = **0.05**)



两个总体均值之差的Z检验
（计算结果）

- H_0: $\mu_1 - \mu_2 = 0$
- H_1: $\mu_1 - \mu_2 \neq 0$
- $\alpha = 0.05$
- $n_1 = 32$，$n_2 = 40$
- 临界值(s):



检验统计量:

$$z = \frac{(\bar{x}_1 - \bar{x}_2) - (\mu_1 - \mu_2)}{\sqrt{\frac{\sigma_1^2}{n_1} + \frac{\sigma_2^2}{n_2}}} = \frac{50 - 40 - 0}{\sqrt{\frac{64}{32} + \frac{100}{40}}} = 2.83$$

决策:

拒绝H_0

结论:

有证据表明两种方法生产的产品其抗拉强度有显著差异

两个总体均值之差的 *t* 检验
(σ_1^2、σ_2^2未知)

1. 检验具有等方差的两个总体的均值
2. 假定条件
   - 两个样本是独立的随机样本
   - 两个总体都是正态分布
   - 两个总体方差未知但相等$\sigma_1^2 = \sigma_2^2$
3. 检验统计量

$$t = \frac{(\bar{x}_1 - \bar{x}_2) - (\mu_1 - \mu_2)}{S_p\sqrt{\frac{1}{n_1} + \frac{1}{n_2}}} \qquad S_p^2 = \frac{(n_1 - 1)S_1^2 + (n_2 - 1)S_2^2}{(n_1 - 1) + (n_2 - 1)}$$

两个总体均值之差的 *t* 检验
(σ_1^2、σ_2^2未知)

1. 检验方差不等的两个总体的均值
2. 假定条件
   - 两个样本是独立的随机样本
   - 两个总体都是正态分布
   - 两个总体方差未知且不相等$\sigma_1^2 \neq \sigma_2^2$
3. 检验统计量

$$t = \frac{(\bar{x}_1 - \bar{x}_2) - (\mu_1 - \mu_2)}{\sqrt{\frac{S_1^2}{n_1} + \frac{S_2^2}{n_2}}} \sim t(df)$$

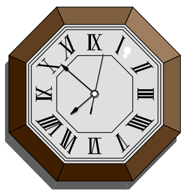
两个总体均值之差的 *t* 检验

【例】研究两种激素类药物对肾组织切片的氧消耗的影响，结果是

（1）$n_1 = 9, \bar{x}_1 = 27.92, S_1^2 = 8.673$;

（2）$n_2 = 6, \bar{x}_2 = 25.11, S_2^2 = 1.843$;

问两种药物对肾组织切片氧消耗的影响差异是否显著？（已知肾组织切片氧消耗量是服从正态分布的随机变量）



(1) 方差齐性检验

H_0: $\sigma_1 = \sigma_2$

H_1: $\sigma_1 \neq \sigma_2$

$\alpha = 0.05$

$$F = \frac{S_1^2}{S_2^2} = \frac{8.673}{1.843} = 4.71$$

$F_{0.025}(8,5) = 6.757$

$\because F < F_{0.025}(8,5)$

在显著性0.05的水平上，可以接受$\sigma_1 = \sigma_2$的假设

(2) 均值之差的显著性检验

H_0: $\mu_1 = \mu_2$

H_1: $\mu_1 \neq \mu_2$

$\alpha = 0.05$

$$t = \frac{\bar{x}_1 - \bar{x}_2}{\sqrt{S_p^2(\frac{1}{n_1} + \frac{1}{n_2})}} = 2.168$$

$t_{0.025} = 2.160$

$\because t > t_{0.025}$

在显著性0.05的水平上，两种药物对肾组织切片氧消耗的影响显著。

两个总体均值之差的 *t* 检验

【例】两组类似的大鼠，一组做对照，另一组做催产素处理，然后测定血糖含量(mg)，结果如下：

对照组：$n_1 = 12$, $\bar{x}_1 = 109.17$, $S_1^2 = 97.430$;

实验组：$n_2 = 8$, $\bar{x}_2 = 106.88$, $S_2^2 = 7.268$;

问药物对大鼠血糖含量的影响是否显著？（已知血糖含量是服从正态分布的随机变量）

(1) 方差齐性检验

H_0: $\sigma_1 = \sigma_2$

H_1: $\sigma_1 \neq \sigma_2$

$\alpha = 0.05$

$$F = \frac{S_1^2}{S_2^2} = 13.41$$

$F_{0.025}(11,7) = 4.714$

$\because F > F_{0.025}(11,7)$

$\therefore$ 两个总体方差不相等

(2) 均值之差的显著性检验

H_0: $\mu_1 = \mu_2$ H_1: $\mu_1 \neq \mu_2$

$\alpha = 0.05$

$$t = \frac{\bar{x}_1 - \bar{x}_2}{\sqrt{S_1^2/n_1 + S_2^2/n_2}} = 0.76$$

$df = 13.35$ $t_{0.025}(13) = 2.16$

$t_{0.025}(14) = 2.145$

$\therefore t < t_{0.025}(df)$

∴在显著性0.05的水平上，催产素对大鼠血糖含量的影响不显著。

假设检验中相关样本的应用

两个相关（配对或匹配）样本的均值检验

两个总体均值之差的检验

（配对样本的 *t* 检验）

1. 检验两个相关总体的均值
   - 配对或匹配
   - 重复测量 (前/后)
2. 利用相关样本可消除项目间的方差
3. 假定条件
   - 两个总体都服从正态分布
   - 如果不服从正态分布，可用正态分布来近似 ($n_1 \geq 30$, $n_2 \geq 30$)

配对样本的 *t* 检验

（数据形式）

观察序号	样本1	样本2	差值
1	x_{11}	x_{21}	$D_1 = x_{11} - x_{21}$
2	x_{12}	x_{22}	$D_1 = x_{12} - x_{22}$
M	M	M	M
i	x_{1i}	x_{2i}	$D_1 = x_{1i} - x_{2i}$
M	M	M	M
n	x_{1n}	x_{2n}	$D_1 = x_{1n} - x_{2n}$

配对样本的 *t* 检验

(假设的形式)

假设	研究的问题		
	没有差异 有差异	总体$_1$ ≥ 总体$_2$ 总体$_1$ < 总体$_2$	总体$_1$ ≥ 总体$_2$ 总体$_1$ > 总体$_2$
H_0	$\mu_D = 0$	$\mu_D \geq 0$	$\mu_D \leq 0$
H_1	$\mu_D \neq 0$	$\mu_D < 0$	$\mu_D > 0$

注：$D_i = X_{1i} - X_{2i}$，对第 *i* 对观察值

配对样本的 *t* 检验

（检验统计量）

统计量

$$t = \frac{\bar{x}_D - \mu_D}{S_D/\sqrt{n_D}}$$ 自由度$df = n_D - 1$

样本均值

$$\bar{x}_D = \frac{\sum_{i=1}^{n} D_i}{n_D}$$

样本标准差

$$s_D = \sqrt{\frac{\sum_{i=1}^{n}(D_i - \bar{x}_D)^2}{n_D - 1}}$$

配对样本的 *t* 检验

【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**公斤以上。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 α **= 0.05**的显著性水平下，调查结果是否支持该俱乐部的声称？

属于检验某项声明的假设！

配对样本的 *t* 检验
（计算表）

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

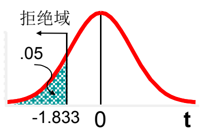
H0: $\mu_1 - \mu_2 \geq 8.5$

H1: $\mu_1 - \mu_2 < 8.5$

α = 0.05

df = 10 - 1 = 9

临界值(s):



$$\bar{x}_D = \frac{\sum_{i=1}^{n} D_i}{n_D} = 9.85$$

$$S_D = \sqrt{\frac{\sum_{i=1}^{n}(D_i - \bar{x}_D)^2}{n_D - 1}} = 2.199$$

$$t = \frac{\bar{x}_D - \mu_D}{S_D/\sqrt{n_D}} = \frac{9.85 - 8.5}{2.199/\sqrt{10}} = 1.941$$

决策: 接受H_0

结论: 可以认为该俱乐部的宣称是可信的

两个总体比例之差的检验
（*Z* 检验）

两个总体比例之差的Z检验

1. 假定条件
   - 两个总体是独立的
   - 两个总体都服从二项分布
   - 可以用正态分布来近似
2. 检验统计量

$$z = \frac{(\hat{P}_1 - \hat{P}_2) - (P_1 - P_2)}{\sqrt{\frac{\hat{P}_1(1-\hat{P}_1)}{n_1} + \frac{\hat{P}_2(1-\hat{P}_2)}{n_2}}} \sim N(0,1)$$

两个总体比例之差的检验
（假设的形式）

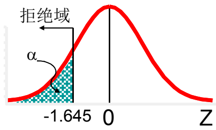
假设	研究的问题		
	没有差异 有差异	比例$_1$ ≥比例$_2$ 比例$_1$ < 比例$_2$	总体$_1$ ≤比例$_2$ 总体$_1$ > 比例$_2$
H_0	$P_1-P_2 = 0$	$P_1-P_2 \geq 0$	$P_1-P_2 \leq 0$
H_1	$P_1-P_2 \neq 0$	$P_1-P_2 < 0$	$P_1-P_2 > 0$

两个总体比例之差的Z检验
(例子)

【例】对两个大型企业青年工人参加技术培训的情况进行调查，调查结果如下：甲厂：调查**60**人，**18**人参加技术培训。乙厂调查**40**人，**14**人参加技术培训。能否根据以上调查结果认为乙厂工人参加技术培训的人数比例高于甲厂？(α = **0.05**)

两个总体比例之差的Z检验
(计算结果)

H0: $P_1 - P_2 \ge 0$
H1: $P_1 - P_2 < 0$
α = 0.05
$n_1 = 60$，$n_2 = 40$
临界值(s):



检验统计量:

$$z=\frac{(\hat{P}_1-\hat{P}_2)-(P_1-P_2)}{\sqrt{\frac{\hat{P}_1(1-\hat{P}_1)}{n_1}+\frac{\hat{P}_2(1-\hat{P}_2)}{n_2}}}=\frac{0.30-.035-0}{\sqrt{\frac{0.30(1-030)}{60}+\frac{0.35(1-0.35)}{40}}}=-0.52$$

决策:

接受H_0

结论:

没有证据表明乙厂工人参加技术培训的人数比例高于甲厂

第四节 假设检验中的其他问题

一. 用置信区间进行检验
二. 利用*P* - 值进行检验

假设检验与参数估计

- 为了验证统计报表的正确性，做了共50人的抽样调查，人均收入的结果：均值871元，标准差21元，问能否证明统计报表中人均收入880元是正确的？（显著性水平为0.05）

已知样本数据：$n=50, \bar{x}=871, s=21$

欲验证 H_0: $\mu=880$, 显著性水平 $\alpha=0.05$

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

用区间估计的方法，如何作出判断？

假设检验与参数估计

- 假设检验

$$t_{-\frac{\alpha}{2}}<t=\frac{\bar{x}-\mu_0}{s/\sqrt{n}}<t_{\frac{\alpha}{2}}$$

$$\mu_0-t_{\frac{\alpha}{2}}\cdot\frac{s}{\sqrt{n}}<\bar{x}<\mu_0+t_{\frac{\alpha}{2}}\cdot\frac{s}{\sqrt{n}}$$

- 参数估计

$$P\left(t_{-\frac{\alpha}{2}}<\frac{\bar{x}-\mu_0}{\frac{s}{\sqrt{n}}}<t_{\frac{\alpha}{2}}\right)=1-\alpha$$

$$\left(\bar{x}-t_{\frac{\alpha}{2}}\cdot\frac{s}{\sqrt{n}},\quad \bar{x}+t_{\frac{\alpha}{2}}\cdot\frac{s}{\sqrt{n}}\right)$$

利用置信区间进行假设检验
(双侧检验)

1. 求出双侧检验均值的置信区间

   σ^2已知时：$\left(\bar{x}-z_{\alpha/2}\frac{\sigma}{\sqrt{n}},\bar{x}+z_{\alpha/2}\frac{\sigma}{\sqrt{n}}\right)$

   σ^2未知时：$\left(\bar{x}-t_{\alpha/2}\frac{s_{n-1}}{\sqrt{n}},\bar{x}+t_{\alpha/2}\frac{s_{n-1}}{\sqrt{n}}\right)$

2. 若总体的假设值μ_0在置信区间外，拒绝H_0

利用置信区间进行假设检验
（左侧检验）

1. 求出单边置信下限

$$\bar{x} - z_{\alpha}\frac{\sigma}{\sqrt{n}} \text{或} \bar{x} - t_{\alpha}\frac{s_{n-1}}{\sqrt{n}}$$

2. 若总体的假设值μ_0小于单边置信下限，拒绝H_0

利用置信区间进行假设检验
（右侧检验）

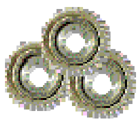
1. 求出单边置信上限

$$\bar{x} + z_{\alpha}\frac{\sigma}{\sqrt{n}} \text{或} \bar{x} + t_{\alpha}\frac{s_{n-1}}{\sqrt{n}}$$

2. 若总体的假设值μ_0大于单边置信上限，拒绝H_0

利用置信区间进行假设检验
(例子)

【例】一种袋装食品每包的标准重量应为**1000**克。现从生产的一批产品中随机抽取**16**袋，测得其平均重量为**991**克。已知这种产品重量服从标准差为**50**克的正态分布。试确定这批产品的包装重量是否合格？(**α = 0.05**)



利用置信区间进行假设检验
（计算结果）

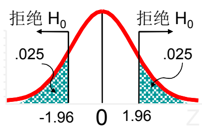
H_0: $\mu = 1000$

H_1: $\mu \neq 1000$

$\alpha = 0.05$

$n = 16$

临界值(s):



置信区间为

$$\left(\bar{x} - z_{\alpha/2}\frac{\sigma}{\sqrt{n}}, \bar{x} + z_{\alpha/2}\frac{\sigma}{\sqrt{n}}\right)$$

$$= \left(991 - 1.96\frac{50}{\sqrt{16}}, 991 + 1.96\frac{50}{\sqrt{16}}\right)$$

$$= (966.5, 1015.5)$$

决策:

假设的μ_0=1000在置信区间内，接受H_0

结论:

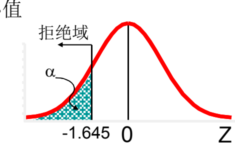
表明这批产品的包装重量合格

观察到的显著性水平
P-值

利用 *P*-值进行假设检验

什么是 *P* 值（*P*-Value）？

1. 是一个概率值
2. 如果我们假设原假设为真，*P*-值是观测到的样本均值不同于(<或 >) 实测值的概率。换句话说，在零假设下，检验统计量取其实现值及（沿着备择假设的方向）更加极端值的概率称为p-值。
3. 被称为观察到的(或实测的)显著性水平
   - H_0 能被拒绝的α的最小值



利用 *P* 值进行决策

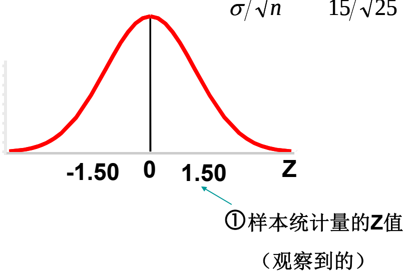
1. 单侧检验
   – 若p-值 ≥ α, 不能拒绝 H_0
   – 若p-值 < α, 拒绝 H_0
2. 双侧检验
   – 若p-值 ≥ $\alpha/2$, 不能拒绝 H_0
   – 若p-值 < $\alpha/2$, 拒绝 H_0

双尾 Z 检验
(*P*-值计算实例)

【例】欣欣儿童食品厂生产的盒装儿童食品每盒的标准重量为368克。现从某天生产的一批食品中随机抽取25盒进行检查，测得每盒的平均重量为 $\bar{x}$ = 372.5克。企业规定每盒重量的标准差σ为15克。确定*P* - 值。

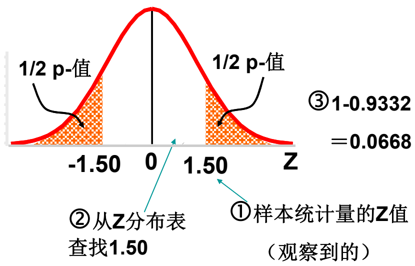
双尾 Z 检验
(*P*-值计算结果)

计算的检验统计量为：$Z=\dfrac{\bar{x}-\mu_0}{\sigma/\sqrt{n}}=\dfrac{372.5-368}{15/\sqrt{25}}=1.5$



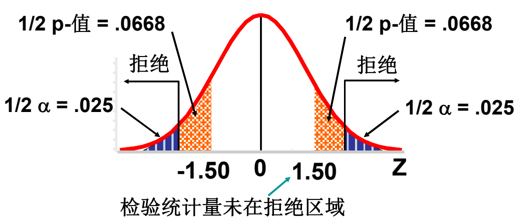
双尾 Z 检验
(*P*-值计算结果)

p-值为 P(Z ≤ -1.50 或 Z ≥ 1.50)



双尾 Z 检验
(*P*-值计算结果)

p = 0.1336 > α = 0.05， 不能拒绝H_0

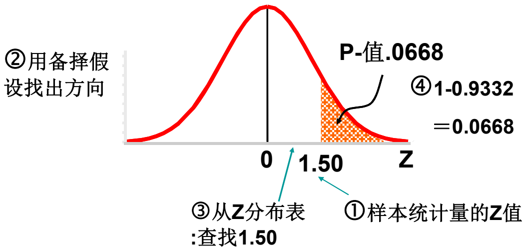


单尾 Z 检验
(*P*-值计算结果)

【例】欣欣儿童食品厂生产的某种盒装儿童食品，规定每盒的重量不高于368克。现从某天生产的一批食品中随机抽取25盒进行检查，测得每盒的平均重量为 $\bar{x}$=372.5克。企业规定每盒重量的标准差σ为15克。确定P-值。

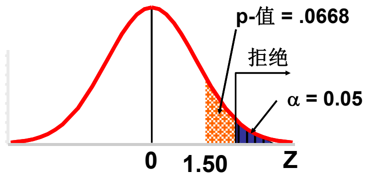
单尾 Z 检验
(*P*-值计算结果)

p-值为 P(Z ≥ 1.50)=.0668



单尾 Z 检验
(*P*-值计算结果)

(p-值 =0 .0668) ≥ (α = 0.05)，不能拒绝H_0



检验统计量未在拒绝区域

本章小结

1. 假设检验的概念和类型
2. 假设检验的过程
3. 基于一个样本的假设检验问题
4. 基于两个样本的假设检验问题
5. 利用置信区间进行假设检验
6. 利用*p* - 值进行假设检验

结 束