

第一讲-3 假设检验

- 1 假设检验的基本原理
 - 2 一个总体参数的检验
 - 3 二个总体参数的检验
 - 4 **SPSS**在参数估计与假设检验中的应用
- 
- A globe is shown with a city skyline, including several skyscrapers and a green tree, on a small patch of land at the top. The globe is tilted and shows a map of the world, with labels for various countries and cities. The background is a blue sky with clouds.

学习目标

- 假设检验的基本思想和原理
- 假设检验的步骤
- 一个总体参数的检验
- ***P***值的计算与应用
- 用**SPSS**进行假设检验



假设检验

1 假设检验的基本原理

- 1.1 怎样提出假设？
- 1.2 怎样做出决策？
- 1.3 怎样表述决策结果？

1 假设检验的基本原理

1.1 怎样提出假设？



假设检验的定义

1. 先对总体的参数(或分布形式)提出某种假设，然后利用样本信息判断假设是否成立的统计方法
2. 逻辑上运用反证法，统计上依据小概率原理
 - 小概率是在一次试验中，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



原假设

1. 又称“0假设”，研究者想收集证据予以反对的假设，用 H_0 表示
2. 所表达的含义总是指参数没有变化或变量之间没有关系
3. 最初被假设是成立的，之后根据样本数据确定是否有足够的证据拒绝它
4. 总是有符号 $=, \leq$ 或 \geq
 - $H_0: \mu =$ 某一数值
 - $H_0: \mu \geq$ 某一数值
 - $H_0: \mu \leq$ 某一数值
 - 例如, $H_0: \mu = 10\text{cm}$



备择假设

1. 也称“研究假设”, H_1 或 H_a 表示
2. 所表达的含义是总体参数发生了变化或变量之间有某种关系
3. 备择假设通常用于表达研究者自己倾向于支持的看法, 然后就是想办法收集证据拒绝原假设, 以支持备择假设
4. 总是有符号 $\neq, <$ 或 $>$
 - $H_1: \mu \neq$ 某一数值
 - $H_1: \mu >$ 某一数值
 - $H_1: \mu <$ 某一数值



双侧检验与单侧检验

1. 备择假设没有特定的方向性，并含有符号“ \neq ”的假设检验，称为双侧检验或双尾检验(**two-tailed test**)
2. 备择假设具有特定的方向性，并含有符号“ $>$ ”或“ $<$ ”的假设检验，称为单侧检验或单尾检验(**one-tailed test**)
 - 备择假设的方向为“ $<$ ”，称为**左侧检验**
 - 备择假设的方向为“ $>$ ”，称为**右侧检验**



提出假设

(例题分析)

❖ **【例】** 一种零件的生产标准是直径应为10cm，为对生产过程进行控制，质量监测人员定期对一台加工机床检查，确定这台机床生产的零件是否符合标准要求。如果零件的平均直径大于或小于10cm，则表明生产过程不正常，必须进行调整。试陈述用来检验生产过程是否正常的原假设和被择假设

解： 研究者想收集证据予以证明的假设应该是“生产过程不正常”。建立的原假设和备择假设为

$$H_0: \mu = 10\text{cm} \quad H_1: \mu \neq 10\text{cm}$$



提出假设

(例题分析)

❖ **【例】** 某品牌洗涤剂在它的产品说明书中声称：平均净含量不少于500克。从消费者的利益出发，有关研究人员要通过抽检其中的一批产品来验证该产品制造商的说明是否属实。试陈述用于检验的原假设与备择假设

解：研究者抽检的意图是倾向于证实这种洗涤剂的平均净含量并不符合说明书中的陈述。建立的原假设和备择假设为

$$H_0: \mu \geq 500 \quad H_1: \mu < 500$$



提出假设

(例题分析)

❖ **【例】** 一家研究机构估计，某城市中家庭拥有汽车的比例超过30%。为验证这一估计是否正确，该研究机构随机抽取了一个样本进行检验。试陈述用于检验的原假设与备择假设。

解： 研究者想收集证据予以支持的假设是“该城市中家庭拥有汽车的比例超过30%”。建立的原假设和备择假设为

$$H_0: \pi \leq 30\% \quad H_1: \pi > 30\%$$



提出假设

(结论与建议)

1. 原假设和备择假设是一个完备事件组，而且相互对立
 - 在一项假设检验中，原假设和备择假设必有一个成立，而且只有一个成立
2. 先确定备择假设，再确定原假设
3. 等号“=”总是放在原假设上
4. 因研究目的不同，对同一问题可能提出不同的假设(也可能得出不同的结论)



1 假设检验的基本原理

1.2 怎样做出决策？



显著性水平 α

- 1.事先确定的用于拒绝原假设 H_0 时所必须的证据
- 2.能够容忍的犯第 I 类错误的最大概率(上限值)
- 3.原假设为真时，拒绝原假设的概率
 - 抽样分布的拒绝域
4. 表示为 α (alpha)
 - 常用的 值有0.01, 0.05, 0.10
5. 由研究者事先确定



依据什么做出决策？

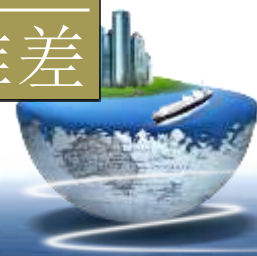
1. 若假设为 $H_0: \mu=500$, $H_1: \mu \neq 500$ 。样本均值为495, 拒绝 H_0 吗？样本均值为502, 拒绝 H_0 吗？
2. 做出拒绝或不拒绝原假设的依据是什么？
3. 传统上, 做出决策所依据的是样本统计量, 现代检验中人们直接使用由统计量算出的犯第 I 类错误的概率, 即所谓的 α 值



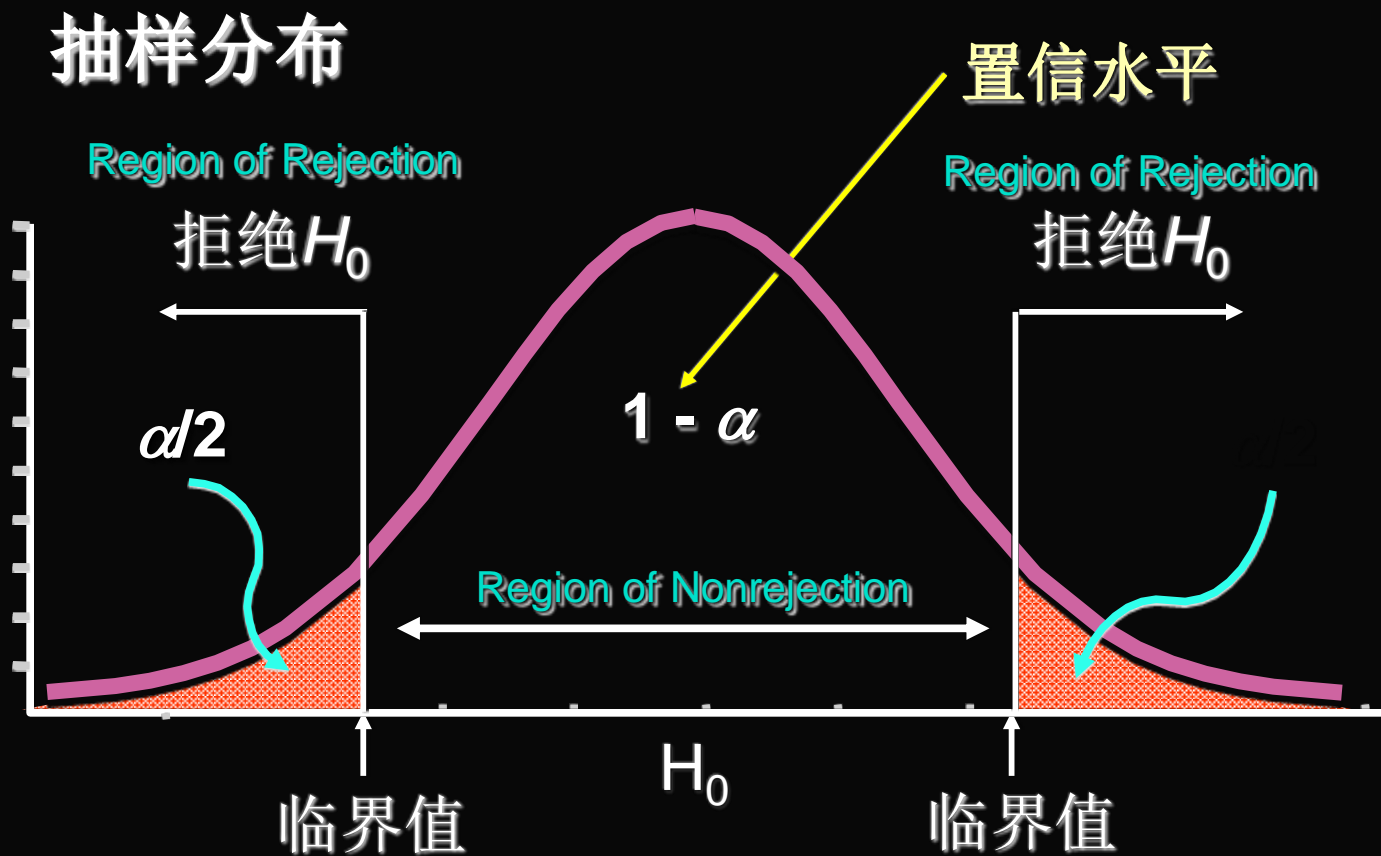
检验统计量

1. 根据样本观测结果计算出对原假设和备择假设做出决策某个样本统计量
2. 对样本估计量的标准化结果
 - 原假设 H_0 为真
 - 点估计量的抽样分布
3. 标准化的检验统计量

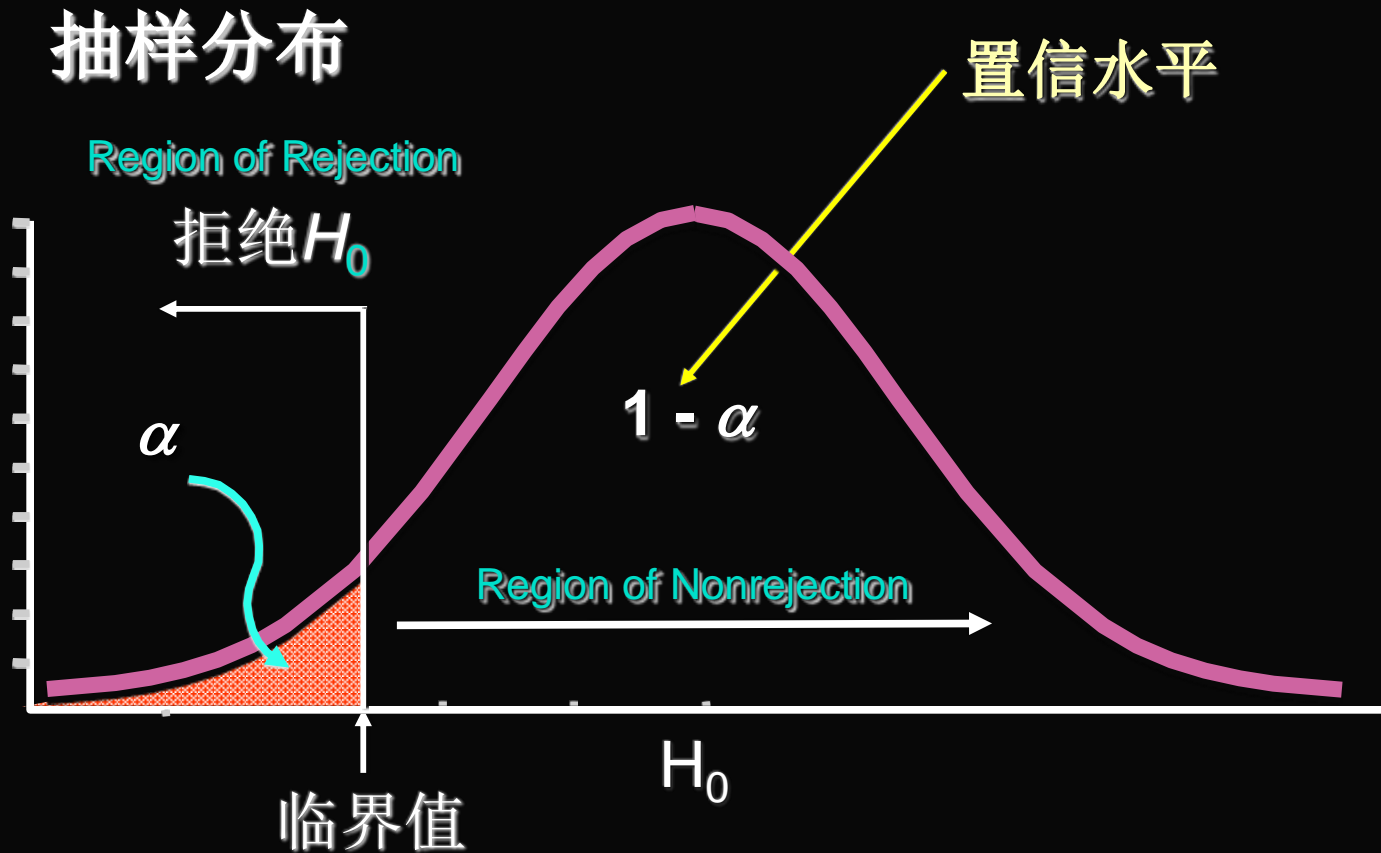
$$\text{标准化检验统计量} = \frac{\text{点估计量} - \text{假设值}}{\text{点估计量的抽样标准差}}$$



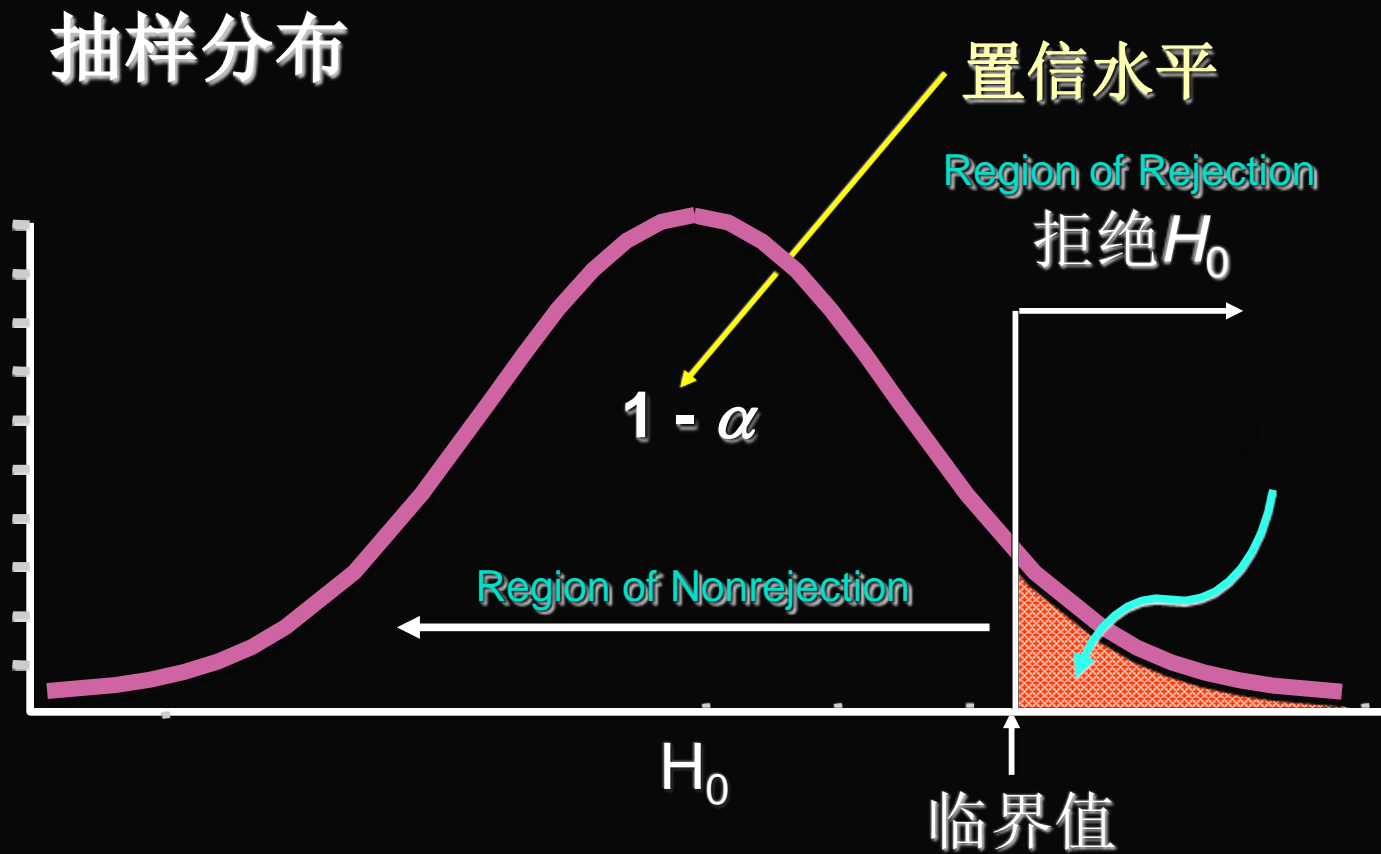
用统计量决策 (双侧检验)



用统计量决策 (左侧检验)



用统计量决策 (右侧检验)



统计量决策规则

1. 给定显著性水平 α ，查表得出相应的临界值 z_{α} 或 $z_{\alpha/2}$ ， t_{α} 或 $t_{\alpha/2}$
2. 将检验统计量的值与 α 水平的临界值进行比较
3. 作出决策
 - 双侧检验： **$|统计量| > 临界值$ ，拒绝 H_0**
 - 左侧检验： **$统计量 < -临界值$ ，拒绝 H_0**
 - 右侧检验： **$统计量 > 临界值$ ，拒绝 H_0**



用P 值决策

1. 如果原假设为真，所得到的样本结果会像实际观测结果那么极端或更极端的概率
 - **P值告诉我们：**如果原假设是正确的话，我们得到得到目前这个样本数据的可能性有多大，如果这个可能性很小，就应该拒绝原假设
2. 被称为观察到的(或实测的)显著性水平
3. 决策规则：若**p值** $< \alpha$, 拒绝 H_0



设 \bar{X}_0 表示抽样得到的 \bar{X} 的观察值，则按 z 变量的变化范围 P 值为

在右单尾检验中， P 值 = $P\left(z \geq \frac{\bar{X}_0 - \mu}{\sigma/\sqrt{n}}\right)$ ；

在左单尾检验中， P 值 = $P\left(z \leq \frac{\bar{X}_0 - \mu}{\sigma/\sqrt{n}}\right)$ ；

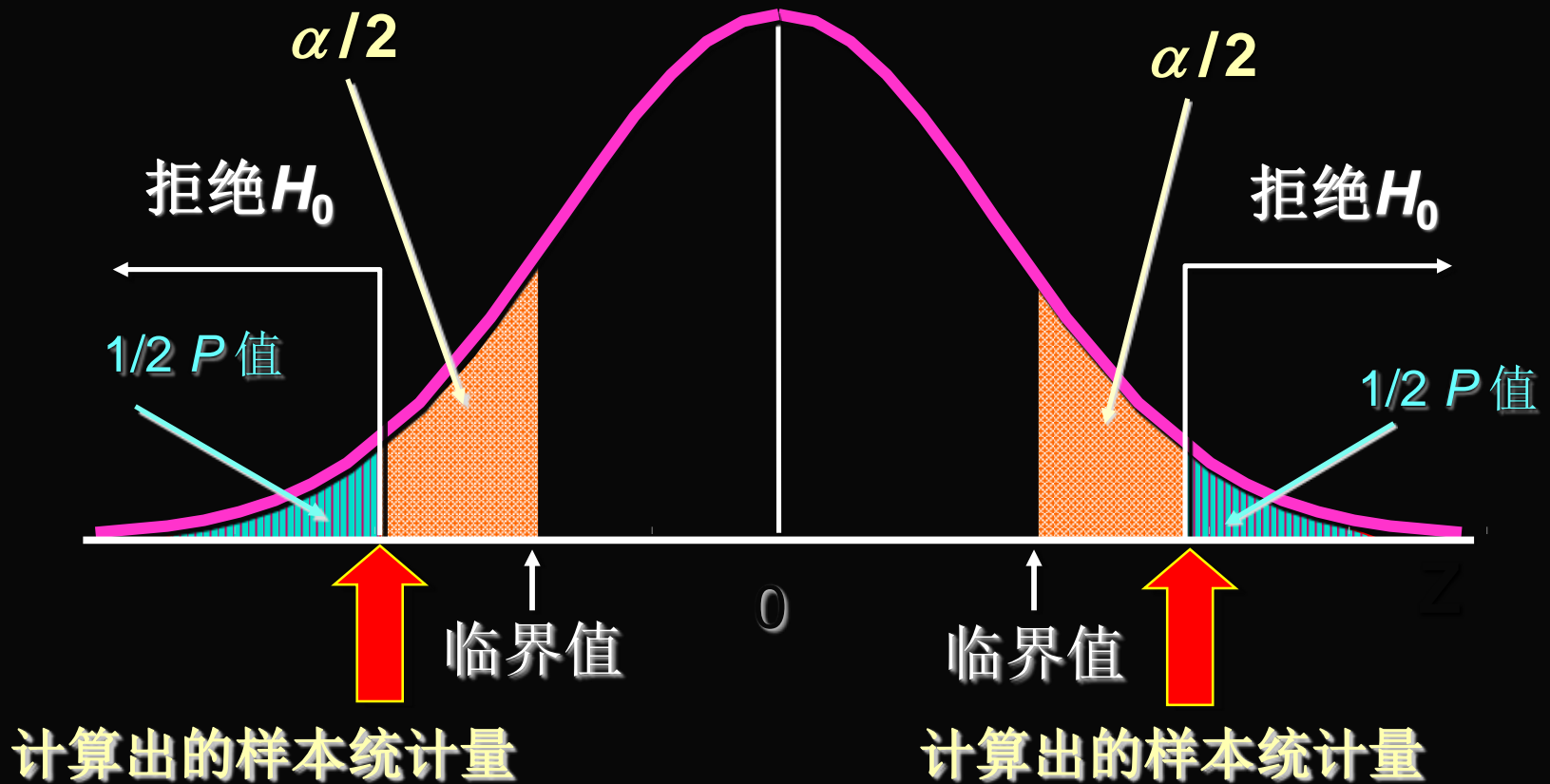
在双尾检验中，

$$P\text{值} = P\left[Z \geq \left|\frac{\bar{X}_0 - \mu}{\sigma/\sqrt{n}}\right|\right] = 2P\left[Z \geq \frac{\bar{X}_0 - \mu}{\sigma/\sqrt{n}}\right] = 2P\left[Z \leq \frac{\bar{X}_0 - \mu}{\sigma/\sqrt{n}}\right]$$

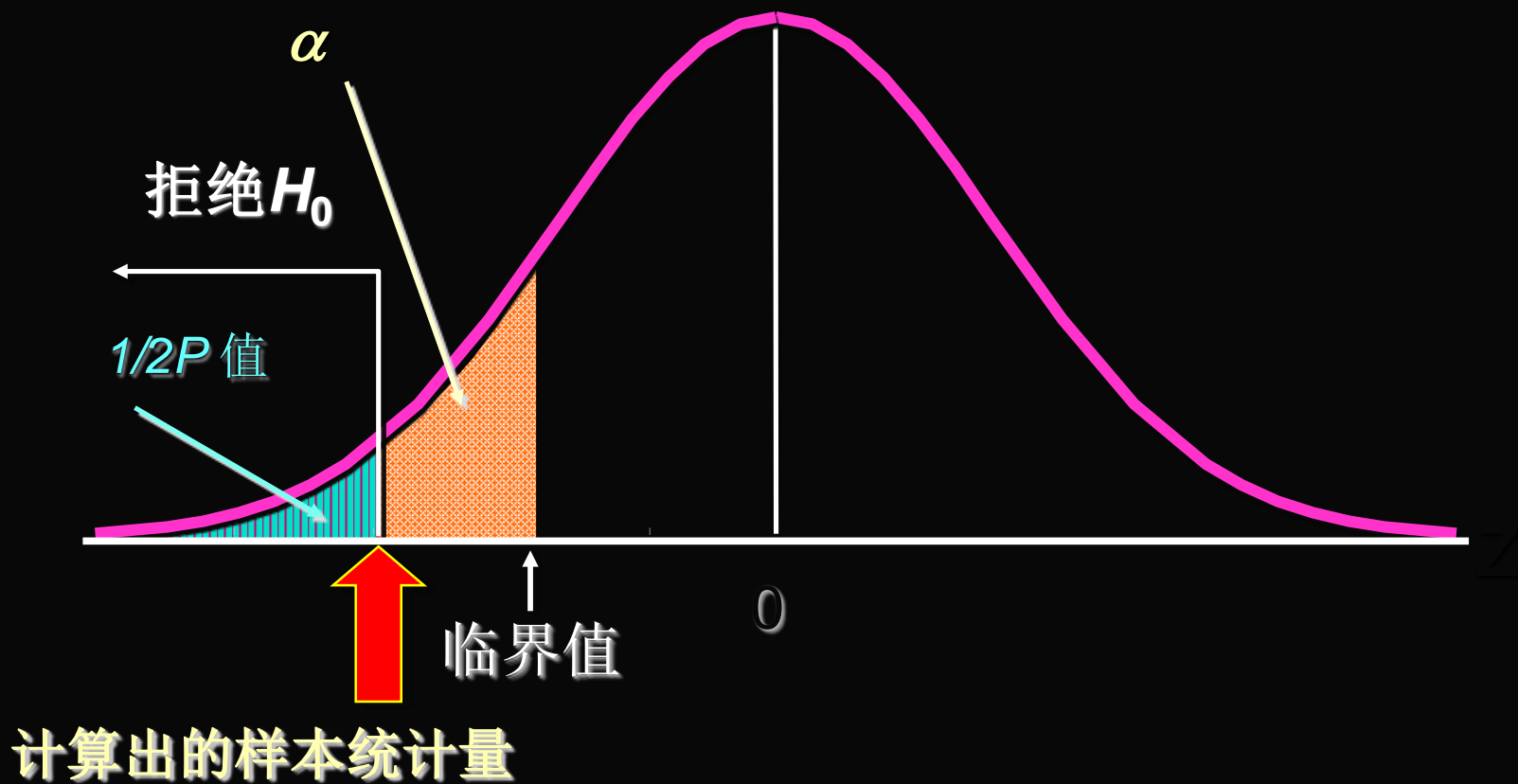
使用 P 值与 α 比较进行判断，显然对于显著程度认识的更加明确。



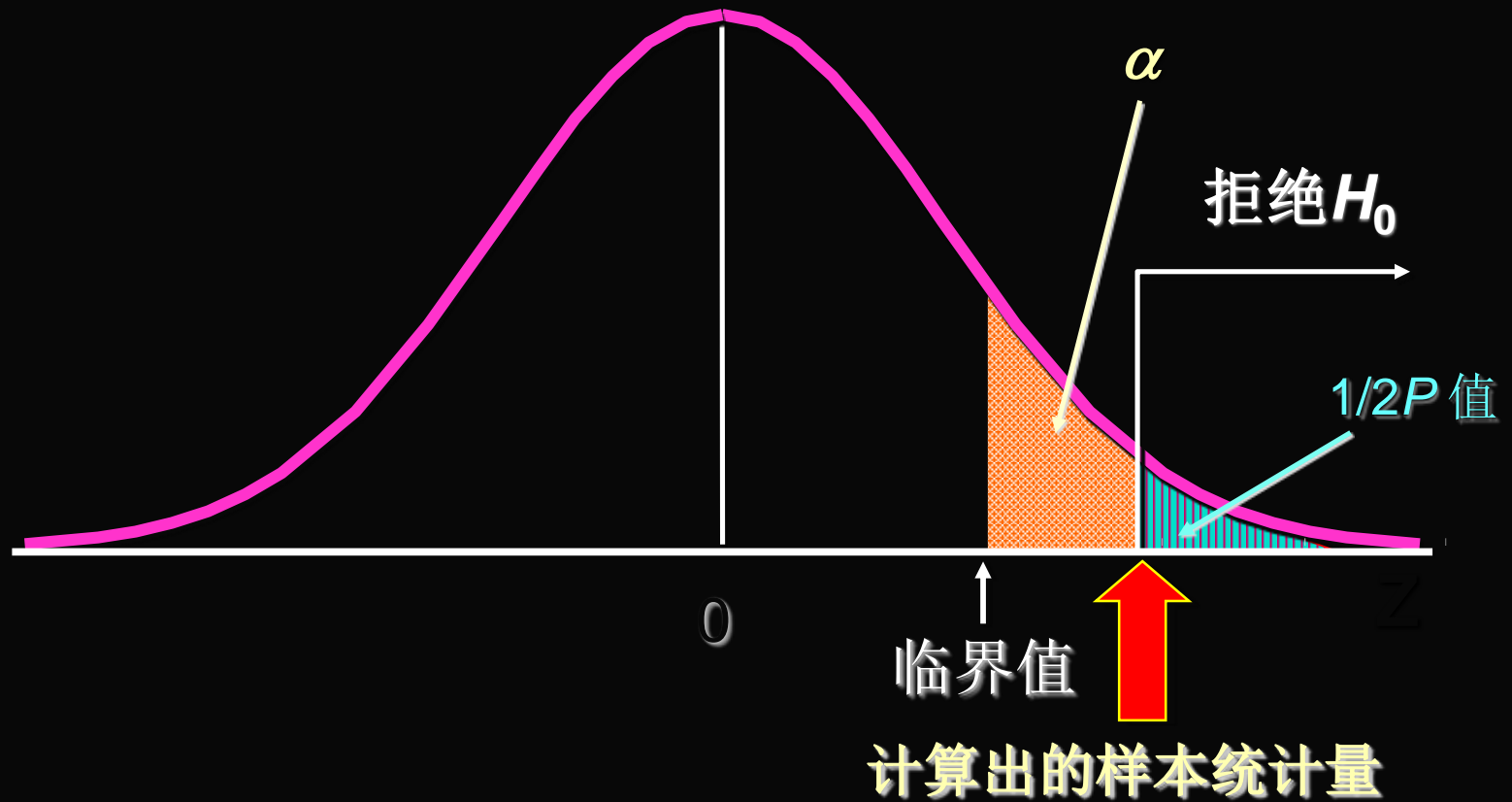
双侧检验的P 值



左侧检验的P 值



右侧检验的P 值



1 假设检验的基本原理

1.3 怎样表述决策结果?



假设检验的基本程序（步骤）

❖ 第一步：

确定零假设和备择假设；选定显著性水平；抽取样本容量为 n 的样本

❖ 第二步：

确定统计量的抽样分布；计算相应的统计量

❖ 第三步：

计算临界值

❖ 第四步：

确定决策规则

❖ 第五步：

判断是否接受零假设；得出结论



假设检验

2 一个总体参数的检验

2.1 总体均值的检验



2 一个总体参数的检验

2.1 总体均值的检验



总体均值的检验

❖ 【例】一种机床加工的零件尺寸绝对平均误差为1.35mm。生产厂家现采用一种新的机床进行加工以期进一步降低误差。为检验新机床加工的零件平均误差与旧机床相比是否有显著降低，从某天生产的零件中随机抽取50个进行检验。利用这些样本数据，检验新机床加工的零件尺寸的平均误差与旧机床相比是否有显著降低？($\alpha=0.01$)

50个零件尺寸的误差数据 (mm)

1.26	1.19	1.31	0.97	1.81
1.13	0.96	1.06	1.00	0.94
0.98	1.10	1.12	1.03	1.16
1.12	1.12	0.95	1.02	1.13
1.23	0.74	1.50	0.50	0.59
0.99	1.45	1.24	1.01	2.03
1.98	1.97	0.91	1.22	1.06
1.11	1.54	1.08	1.10	1.64
1.70	2.37	1.38	1.60	1.26
1.17	1.12	1.23	0.82	0.86

左侧检验

总体均值的检验

(例题分析—大样本)

❖ $H_0 : \mu \geq 1.35$

❖ $H_1 : \mu < 1.35$

❖ $\alpha = 0.01$

❖ $n = 50$

❖ 临界值(c):

检验统计量:

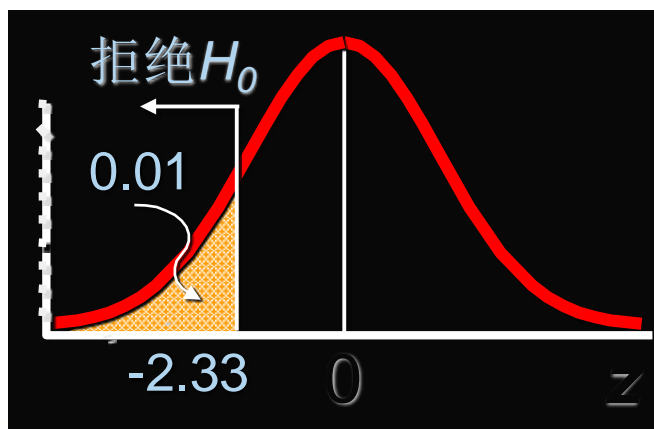
$$z = \frac{1.3152 - 1.35}{0.365749 / \sqrt{50}} = -2.6061$$

决策:

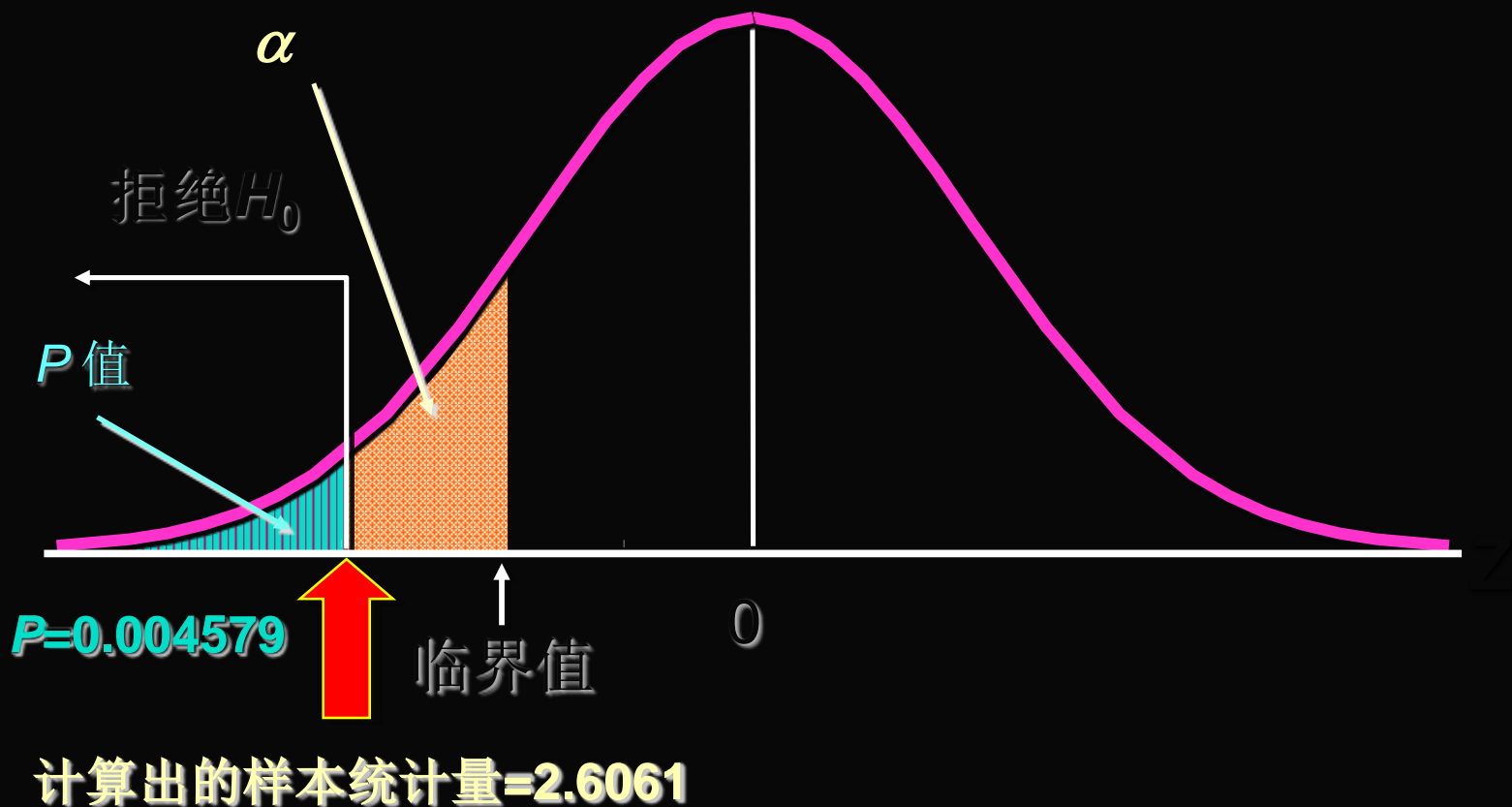
拒绝 H_0

结论:

新机床加工的零件尺寸的平均误差与旧机床相比有显著降低



总体均值的检验 (P 值的图示)



假设检验

3 两个总体参数的检验

3.1 两个总体均值之差的检验

3.2 配对问题的检验

3 两个总体参数的检验

3.1 两个总体均值之差的检验



两个总体均值之差的检验

(独立大样本)

❖ 1. 假定条件

- 两个样本是独立的随机样本
- 正态总体或非正态总体大样本($n_1 \geq 30$ 和 $n_2 \geq 30$)

2. 检验统计量

- σ_1^2, σ_2^2 已知

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

- σ_1^2, σ_2^2 未知

$$z = \frac{(\bar{x}_1 - \bar{x}_2) - (\mu_1 - \mu_2)}{\sqrt{\frac{s_1^2}{n_1} + \frac{s_2^2}{n_2}}} \sim N(0,1)$$



两个总体均值之差的检验

(例题分析—独立大样本)

【例】某公司对男女职员平均小时工资进行了调查，独立抽取了具有同类工作经验的男女职员的两个随机样本，并记录下两个样本的均值、方差等资料如右表。在显著性水平为0.05的条件下，能否认为男性职员与女性职员的平均小时工资存在显著差异？

两个样本的有关数据	
男性职员	女性职员
$n_1=44$	$n_2=32$
$\bar{x}_1=75$	$\bar{x}_2=70$
$S_1^2=64$	$S_2^2=42.25$



两个总体均值之差的检验

(例题分析—独立大样本)

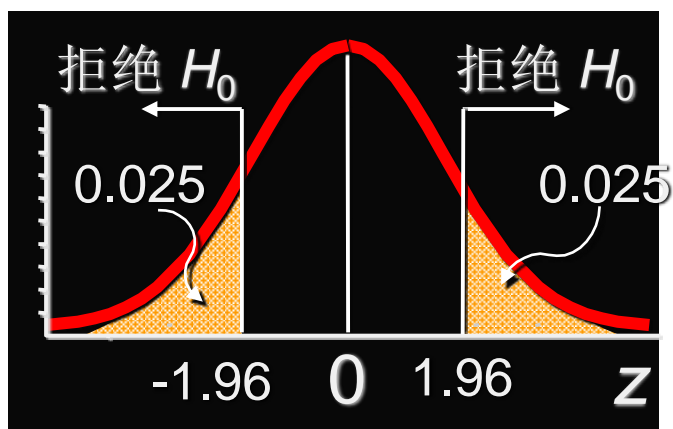
❖ $H_0: \mu_1 - \mu_2 = 0$

❖ $H_1: \mu_1 - \mu_2 \neq 0$

❖ $\alpha = 0.05$

❖ $n_1 = 44, n_2 = 32$

❖ 临界值(c):



检验统计量:

$$z = \frac{75 - 70}{\sqrt{\frac{64}{44} + \frac{42.25}{32}}} = 3.002$$

决策:

拒绝 H_0

结论:

该公司男女职员的平均小时工资之间存在显著差异



3.2 两个总体均值之差的检验 (配对样本)

1. 假定条件

- 两个总体配对差值构成的总体服从正态分布
- 配对差是由差值总体中随机抽取的
- 数据配对或匹配(重复测量 (前/后))


2. 检验统计量

$$t = \frac{\bar{d} - d_0}{s_d / \sqrt{n_d}} \sim t(n-1)$$

样本差值均值

样本差值标准差

$$\bar{d} = \frac{\sum_{i=1}^n d_i}{n_d}$$

$$s_d = \sqrt{\frac{\sum_{i=1}^n (d_i - \bar{d})^2}{n_d - 1}}$$


匹配样本 (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}	$d_2 = x_{12} - x_{22}$
\vdots	\vdots	\vdots	\vdots
i	x_{1i}	x_{2i}	$d_i = x_{1i} - x_{2i}$
\vdots	\vdots	\vdots	\vdots
n	x_{1n}	x_{2n}	$d_n = x_{1n} - x_{2n}$

两总体平均数间差异检验 (配对样本)

员工培训前后操作差错数目对比

	训前差错	训后差错
1	79	65
2	65	64
3	60	50
4	52	48
5	80	72
6	76	69
7	48	45
8	90	80

Paired Samples Statistics

		Mean	N	Std. Deviation	Std. Error Mean
Pair 1	训前	68.75	8	14.82	5.24
	训后	61.63	8	12.61	4.46

Paired Samples Test

	Paired Differences					t	df	Sig. (2-tailed)
	Mean	Std. Deviation	Std. Error Mean	95% Confidence Interval of the Difference				
				Lower	Upper			
Pair 1 前测成绩 - 后测成绩	7.13	4.29	1.52	3.54	10.71	4.697	7	.002

两个总体均值之差的检验

(例题分析—配对样本)

【例】某饮料公司开发研制出一新产品，为比较消费者对新老产品口感的满意程度，该公司随机抽选一组消费者(8人)，每个消费者先品尝一种饮料，然后再品尝另一种饮料，两种饮料的品尝顺序是随机的，而后每个消费者要对两种饮料分别进行评分(0分~10分)，评分结果如下表。取显著性水平 $\alpha = 0.05$ ，该公司是否有证据认为消费者对两种饮料的评分存在显著差异？

两种饮料平均等级的样本数据

旧饮料	5	4	7	3	5	8	5	6
新饮料	6	6	7	4	3	9	7	6



配对总体均值之差的检验

(用Excel进行检验)



	A	B	C
1	t-检验：成对双样本均值分析		
2			
3		变量 1	变量 2
4	平均	5.375	6
5	方差	2.553571429	3.428571429
6	观测值	8	8
7	柏松相关系数	0.724206824	
8	假设平均差	0	
9	df	7	
10	t Stat	-1.357241785	
11	P(T<=t) 单尾	0.108418773	
12	t 单尾临界	1.894578604	
13	P(T<=t) 双尾	0.216837546	
14	t 双尾临界	2.364624251	

案例分析

---SPSS在参数估计与假设检验中的应用

❖ 1.SPSS在参数估计中的应用

❖ 2.SPSS在假设检验中的应用



Thank You !