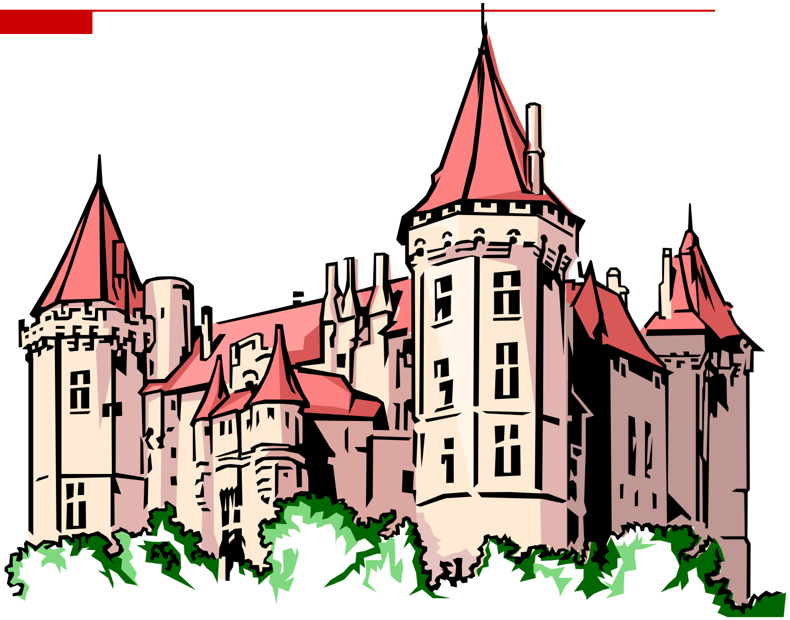
4 推断统计



4.1 引例

例4.1.1：CJW公司客户满意度区间估计

- CJW有限公司十分重视顾客服务的调查研究。近期，CJW的市场部为评估所有CJW顾客满意程度展开了一次抽样调查，随机抽取了100名顾客进行调查，记录他们的满意得分（单位：分），如下表所示：

表4.1.1 100名顾客满意得分

98	86	63	62	48	25	59	92	94	95
79	96	80	84	45	20	97	28	98	96
100	74	76	95	100	94	81	47	97	77
86	51	98	95	94	53	97	95	99	84
89	88	97	99	71	78	98	54	100	44
55	100	64	76	96	90	95	98	95	61
68	100	100	100	90	98	98	75	79	96
37	100	100	96	97	15	62	99	100	96
63	95	98	78	80	97	96	49	95	97
99	100	98	51	89	97	97	94	69	66

4.1 引例（续）

现在，CJW市场部经理要求你根据上述数据完成以下统计分析：

（1）根据以往多次的该类调查发现，CJW顾客的满意得分的标准差在20左右，那么，在95%的置信水平下，CJW顾客的满意得分的置信区间为多少？

（2）若经理认为，以往得到的顾客满意得分的标准差不准确，而对这次调查结果比较满意，那么，在95%置信水平下，你又该如何根据这次调查数据估计CJW顾客满意得分的？

（3）若公司总经理认为，CJW公司顾客的满意得分肯定在80分以上，请你根据这些调查数据在5%的显著性水平下验证总经理的推断。

4.1 引例（续）

例4.1.2：某公司的年末审计时间

- 某公司随机抽取了20个企业进行年末审计，并准确记录了工作人员圆满完成审计任务的时间（单位：小时），如下表所示。总经理为了解公司平均审计时间，请你完成如下问题：

表4.1.2 20次年末审计时间

12	14	21	22	17
15	14	18	33	23
20	15	19	16	28
22	27	18	18	13

（1）在95%置信水平下，根据这些数据估计公司平均年末审计时间；

（2）在5%显著性水平下，验证公司平均年末审计时间是否等于过去估计的**18**小时。

4.1 引例（续）

□ 推断统计分析过程

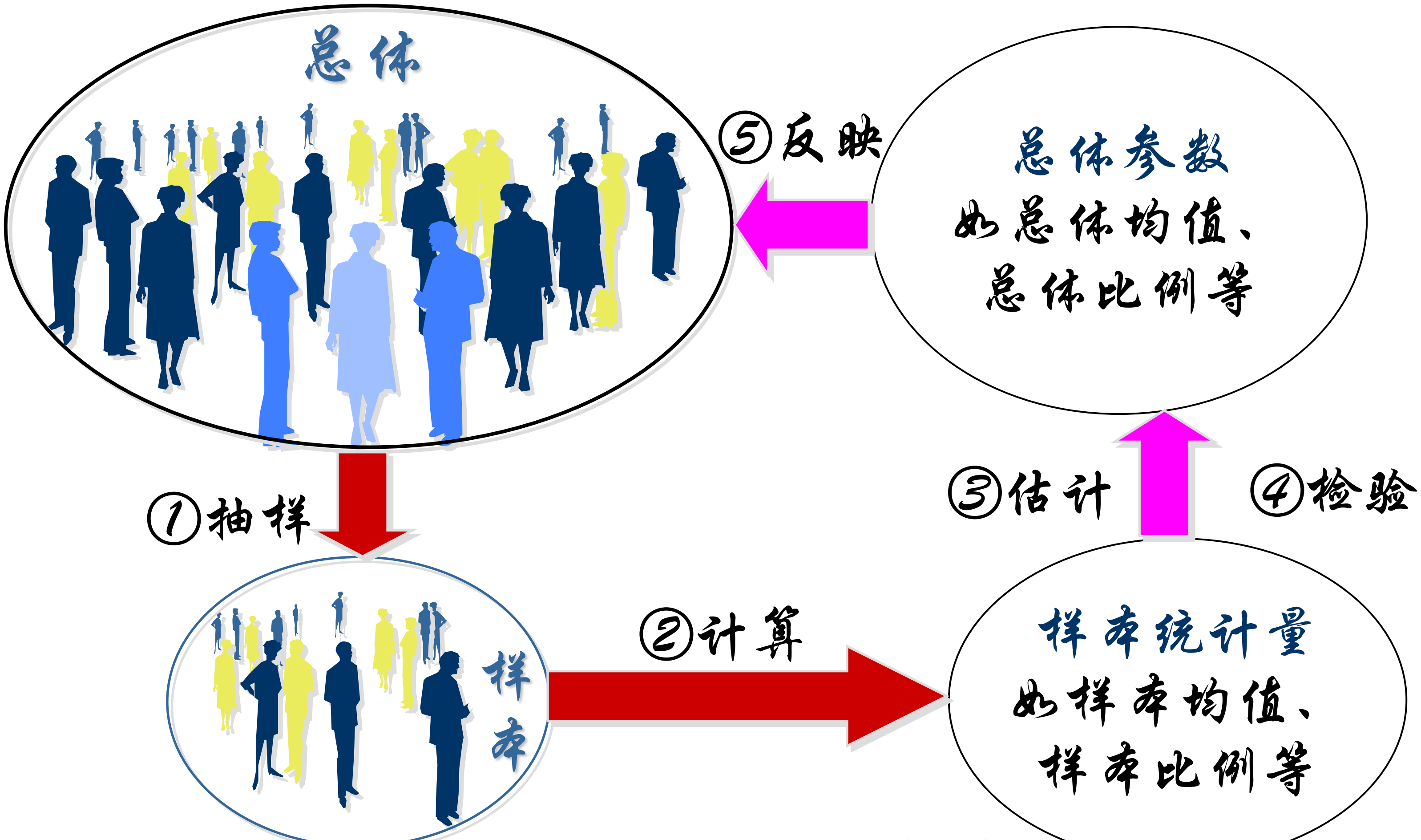


图4.1.1 推断统计分析过程

4.2 参数估计

▫ 参数估计过程

- 参数估计包含**点估计**和**区间估计**两方面。

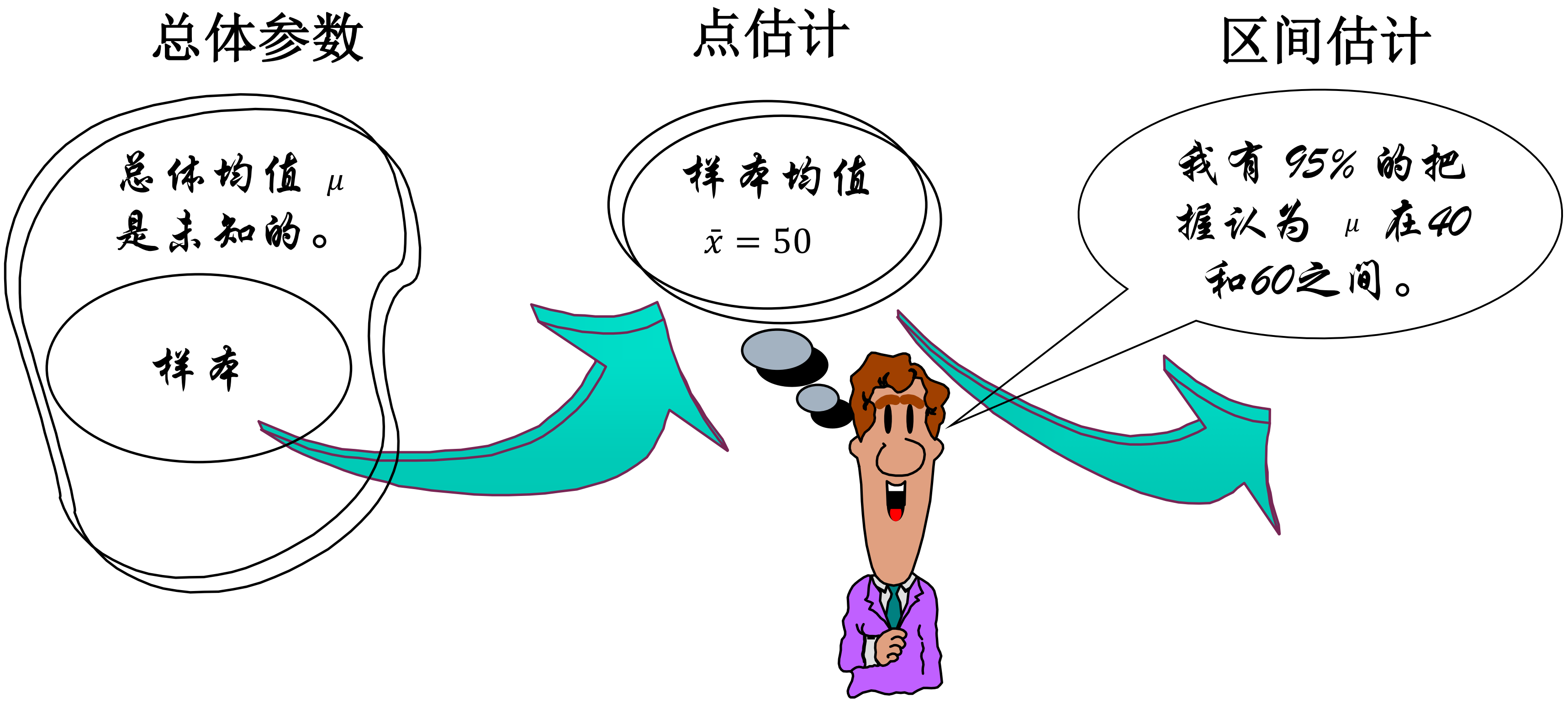


图4.2.1 参数估计过程

4.2 参数估计（续）

点估计（Point Estimation）

- 定义
  - 直接将样本统计量计算值作为总体参数估计值。比如，样本均值的计算值为总体均值的估计值。
- 优良标准：衡量样本统计量估计总体参数的精确程度。
  - 无偏性（Unbiasedness）：样本统计量的均值等于总体参数的真实值。
  - 有效性（Relative Efficiency）：样本统计量的方差越小，其有效性就越大。
  - 一致性（Consistency）：随着样本容量的增加，样本统计量的取值趋近于总体参数真实值的概率逐渐增加（当样本容量趋近于总体单位个数时，样本统计量取值趋近于总体参数真实值的概率接近于1）。

4.2 参数估计（续）

点估计的误差

- **总误差定义**：点估计值与总体参数真实值之差，即 $\Delta = |\hat{\theta} - \theta|$。
- 总误差构成：

（1）**登记性误差**：由于调查过程中各有关环节工作失误造成的，主要原因可能是：填报、抄录、汇总造成的错误。该类误差在任何调查形式中都可能发生。这类误差属于非抽样误差。

（2）**抽样误差**：由于用样本推断总体时样本的**样本代表性不足**而产生的误差。该类误差可以计算和控制，只有在抽样调查中才会产生。在没有特别说明情况下，估计误差仅考虑抽样误差。

4.2 参数估计（续）

区间估计（Interval Estimation）

定义

- **总体参数的点估计值 $\hat{\theta}$ 减去和加上一个误差 Δ 而形成的区间 $[\hat{\theta}-\Delta, \hat{\theta}+\Delta]$ 对总体参数 θ 做出的估计。**
- 区间估计由以下基本要素组成：

（1）置信区间 $[\hat{\theta}-\Delta, \hat{\theta}+\Delta]$：对总体参数的估计范围，其中，$\hat{\theta}-\Delta$ 为置信下限，$\hat{\theta}+\Delta$ 为置信上限；

（2）置信水平 $1-\alpha$ ：置信区间包含总体参数的概率，或者用置信区间估计总体参数达到一定精度的可靠程度，通常取为99%、95%、90%等。

（3）边际误差 Δ：置信上限与置信下限之差的一半，即为**在置信水平 $1-\alpha$ 条件下的抽样误差（$|\hat{\theta}-\theta|$）的最大值**。

4.2 参数估计（续）

➢ 区间估计图示

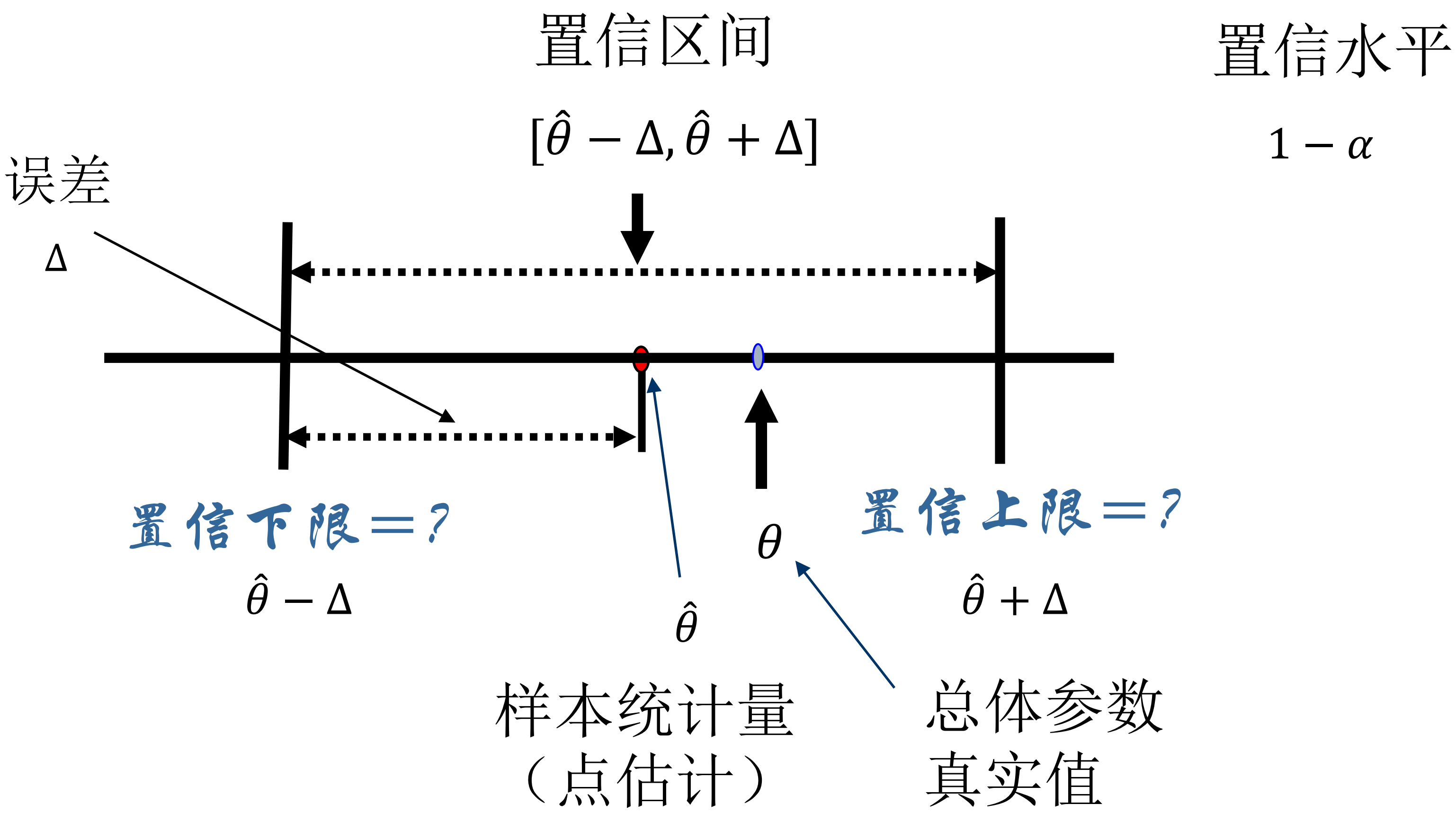


图4.2.2 区间估计

4.2 参数估计（续）

- **区间估计的实际意义**
  - 我们有 $1-\alpha$ 的把握认为区间估计包含了总体均值；
  - 所有可能的样本均值中有 $1-\alpha$ 的样本均值能够构建一个包含总体均值在内的区间。图示
  - **问题4.2.1**：是否意味着总体均值有 $1-\alpha$ 的可能性落在置信区间内？

- **区间估计基本步骤**
  - 第1步，计算总体参数的点估计值；
  - 第2步，给定置信水平，计算出边际误差；
  - 第3步，计算出置信区间并做出解释和判断。

4.2 参数估计（续）

➢ 总体均值的区间估计

- 当样本为**大样本**时，又分两种情况：

（1）**当总体方差 σ_x^2 已知时**，总体均值 μ 的 $1-\alpha$ 置信区间为

$$\left[\bar{x} - Z_{\alpha/2}\frac{\sigma_{\mathrm{x}}}{\sqrt{n}},\ \ \bar{x} + Z_{\alpha/2}\frac{\sigma_{\mathrm{x}}}{\sqrt{n}}\right]$$

（2）**当总体方差未知时**，总体方差就用样本方差 s_x^2 来替代，总体均值 μ 的 $1-\alpha$ 置信区间为

$$\left[\bar{x} - Z_{\alpha/2}\frac{\mathrm{s_x}}{\sqrt{n}},\ \ \bar{x} + Z_{\alpha/2}\frac{\mathrm{s_x}}{\sqrt{n}}\right]$$

4.2 参数估计（续）

- 当样本为小样本时，又分三种情况：

（1）**当总体呈正态分布且总体方差已知时**，估计方法与前面大样本、总体方差已知的情况相同。

（2）**当总体呈正态分布，并且总体方差未知时**，总体方差用样本方差来替代，总体均值 μ 的 $1-\alpha$ 置信区间为

$$\left[\bar{x}-t_{\alpha/2}\frac{s}{\sqrt{n}},\ \ \bar{x}+t_{\alpha/2}\frac{s}{\sqrt{n}}\right]$$

其中，t分布的自由度为n-1。

（3）**当总体分布未知时**，要进行区间估计的唯一办法就是扩大样本容量，或使用诸如非参数统计方法等其它方法。

4.2 参数估计（续）

CJW公司客户满意度(例4.1.1)的区间估计

- 问题1
  - 由100个客户的满意度得分可知，

$\sigma = 20$ ，$n = 100$，$\bar{x} = 82$ ，$\alpha = 5\%$ ，$Z_{0.025} = 1.96$ ，

你会查表吗吗?

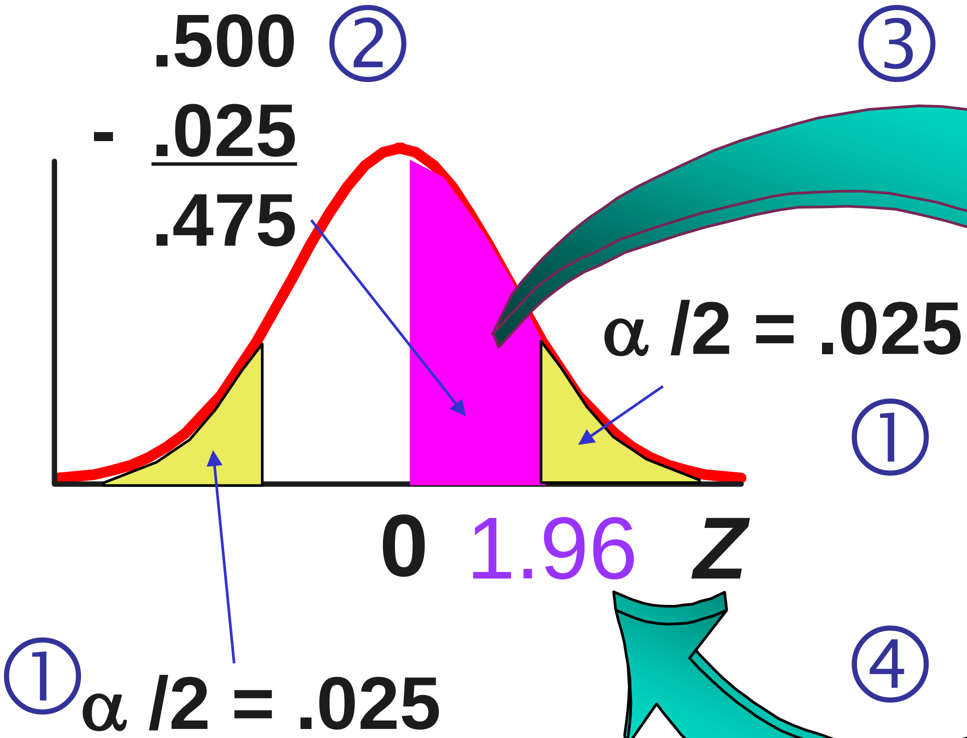
所以，区间估计为：$82 \pm 1.96 \times 20/\sqrt{100}$ ，

即 $[78.08，85.92]$ 。因此，**CJW公司能有95％的把握相信总体平均满意度得分在78.08至85.92间。**

- 问题2
  - 由100个客户满意度得分知，$s = 21$ ，其他同上，所以，区间估计为：$82 \pm 1.96 \times 21/\sqrt{100}$ ，

即 $[77.88，86.12]$ 。因此，**CJW公司能有95％的把握相信总体平均满意度得分在77.88至86.12间。**

当 $\alpha = 0.05$时，$Z_{0.025}$＝?



Z	.05	.06	.07
[illegible]	.4505	.4515	.4525
1.7	.4599	.4608	.4616
1.8	.4678	.4686	.4693
1.9	.4744	.4750	.4756

我会查了，$Z_{0.025}$＝1.96!

4.2 参数估计（续）

□ 某公司年末审计时间(例4.1.2)的区间估计（问题1）

- 由20次年末审计时间计算可得，$\bar{x} = 19.25$，

  $s = 5.44$，$n = 20$，$\alpha = 0.05$，$t_{0.025}(20-1) = 2.093$，

所以，$19.25 \pm 2.093 \times 5.44/\sqrt{20}$， 因此，$[16.70, 21.80]$。

由此，**总经理有95%的把握认为，年末平均审计时间介于16.7至21.8天之间**。

你会查t分布表吗?

重要提示:

当总体方差未知时，均可采用**t**分布进行估计。

4.2 参数估计（续）

- **置信区间宽度的影响因素**
  - **置信水平**：其他因素不变条件下，置信水平越大，置信区间的宽度就越大，估计精度下降。
  - **总体标准差**（或样本标准差）：总体标准差或样本标准差越大，置信区间越宽，估计精度越小。
  - **样本容量**：样本容量增加，置信区间变窄，估计精度提高。
  - 提示1：置信区间估计犹如用网打鱼。
  - 提示2：无偏性、有效性和一致性常用来评价点估计量的优良性，而置信水平和置信区间的宽度常用来评价区间估计。
  - 提示3：区间估计以点估计为基础，但比点估计包含更多的信息（点估计信息+精度信息）。

4.3 假设检验

总体参数假设检验的一般问题

- 什么是假设（Hypothesis）？
  - 关于总体参数值的某种陈述。比如，全国居民月均收入为2654美元。
- 假设类型
  - 零假设（Null Hypothesis），或称原假设（Original Hypothesis）：需要检验是否成立的假设，常用H_0表示，发生的概率常表示为$1-\alpha$；
  - 备择假设（Alternative Hypothesis）：与零假设对立的假设，通常是想要证明的问题，常用H_1表示，发生的概率为α，常称为显著性水平，最常取值为1%、5%、10%。
- 什么是假设检验（Hypothesis Test）？
  - 在样本和概率理论的基础上，对假设是否合理、是否拒绝做出决定的程序。

4.3 假设检验（续）

➢ 假设检验的一般步骤

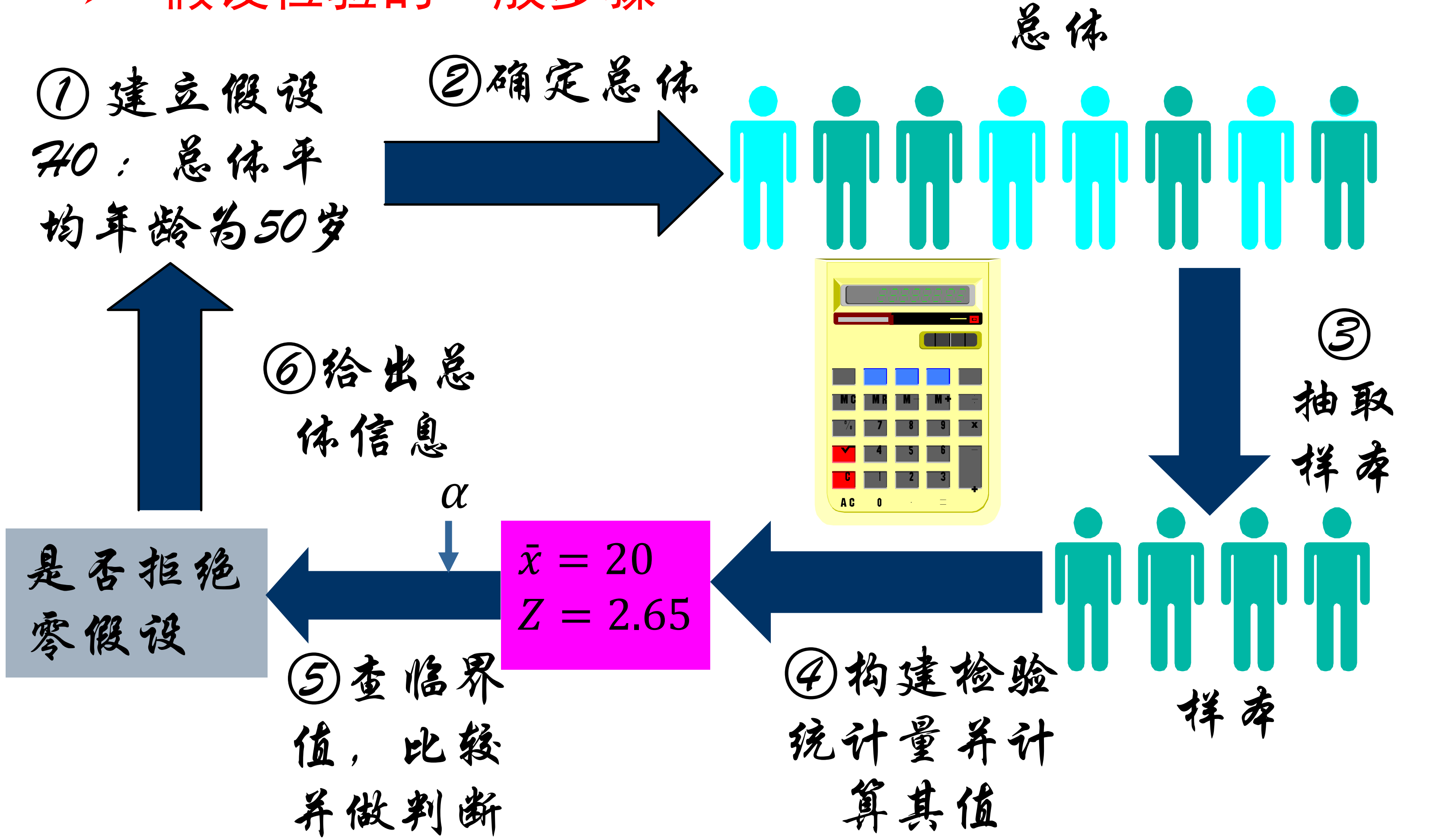


图4.3.1 假设检验的步骤

4.3 假设检验（续）

- **假设检验的类型**

- 左侧检验（Left-tailed Tests）：

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

- 右侧检验（Right-tailed Tests）：

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

- 双侧检验（Two-tailed Tests）：

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

图4.3.2 假设检验类型

4.3 假设检验（续）

总体均值的双侧检验

大样本情况

- 第1步，建立零假设和备择假设：$H_0: \mu = \mu_0$； $H_1: \mu \neq \mu_0$
- 第2步，构建并计算检验统计量

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

- 第3步，查找临界值 $z_{\alpha/2}$，与检验统计量计算值比较，并做出判断。若 $|Z_0| > z_{\alpha/2}$，则拒绝H_0；若 $|Z_0| \leq z_{\alpha/2}$，则不拒绝H_0。

或者，第3步改为：

- 第3步，计算出P值（即 $|Z| > |Z_0|$ 的概率），与显著性水平 α 比较，并做出判断。若 $p \geq \alpha$，则不拒绝H_0；若 $p < \alpha$，则拒绝H_0。

4.3 假设检验（续）

小样本情况

- 当总体分布近似正态分布且总体方差已知时，检验程序与前面的大样本情况完全一样。
- 当总体分布近似正态分布且总体方差未知时，检验统计量：

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

在 α 下，若 $|t|\le t_{\alpha/2}$，不能拒绝H_0，否则拒绝H_0。或者，若 $p\ge\alpha$ ，不能拒绝H_0，否则拒绝H_0。

4.3 假设检验（续）

总体均值的右侧检验

- 大样本情况

  - 第1步，建立零假设和备择假设：

$$H_0: \mu \le \mu_0; \quad H_1: \mu > \mu_0$$

  - 第2步，构建并计算检验统计量

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

  - 第3步，查找临界值 Z_α，与检验统计量计算值比较，并做出判断。若 $Z_0 > Z_\alpha$，则拒绝H_0，否则不能拒绝H_0。

或者，第3步改为：

  - 第3步，计算出P值（即 $Z > Z_0$ 的概率），与显著性水平α 比较，并做出判断。若 $p \ge \alpha$，则不拒绝H_0；若 $p < \alpha$，则拒绝H_0。

总体均值的左侧检验

大样本情况

- 第1步，建立零假设和备择假设：

$$H_0: \mu \geq \mu_0; \quad H_1: \mu < \mu_0$$

- 第2步，构建并计算检验统计量

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

- 第3步，查找临界值 z_α，与检验统计量计算值比较，并做出判断。若 $z_0 < -z_\alpha$ ，则拒绝H_0，否则不能拒绝H_0。

或者，第3步改为：

- 第3步，计算出P值（即 $Z < Z_0$ 的概率），与显著性水平 α 比较，并做出判断。若 $p \geq \alpha$，则不拒绝H0；若 $p < \alpha$，则拒绝H0。

4.3 假设检验（续）

- **小样本情况**
  - 检验统计量及其分布与双侧检验的小样本情况是一样的；
  - 拒绝域或者p值的计算与单侧检验的大样本情况是一样的。
- **提醒**
  - 上述讨论的都是假定总体分布为正态分布。如果总体分布不是正态分布，而且是小样本情况时，需要扩大样本容量或者寻求其他方法！

4.3 假设检验（续）

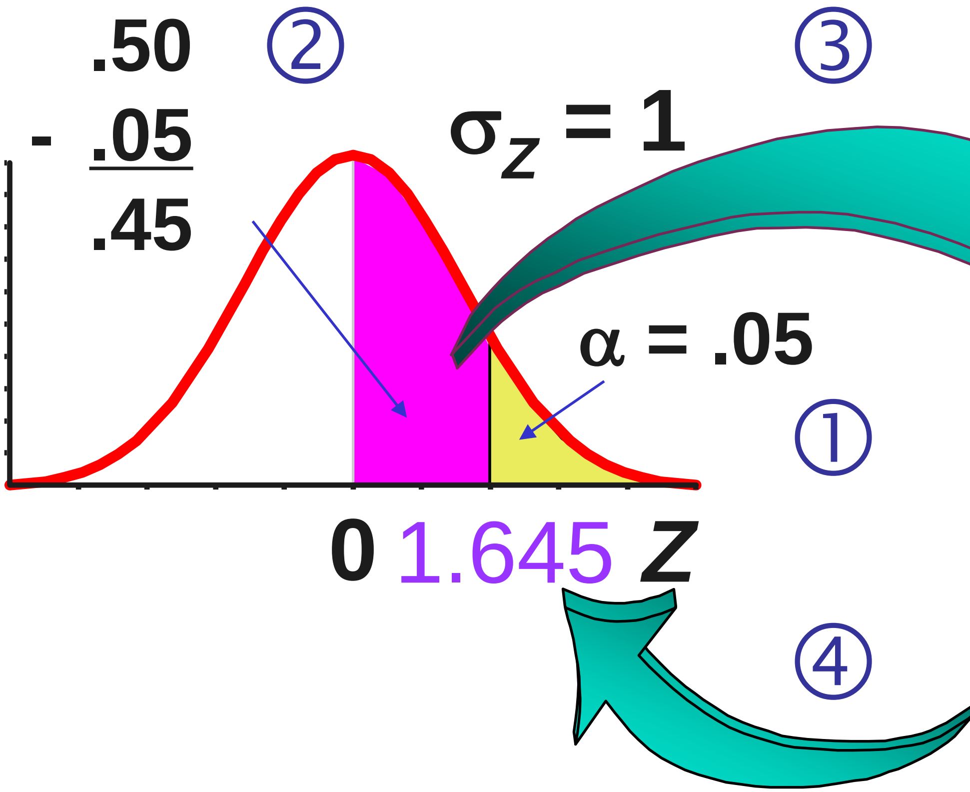
CJW公司客户满意度(例4.1.1)的假设检验（问题3）

- 建立假设： $H_0: \mu \le 80$ ； $H_1: \mu > 80$
- 构建统计量： $Z = \dfrac{\bar{x} - \mu_0}{\sigma/\sqrt{n}} \sim N(0,1)$
- 计算检验统计量：$\bar{x} = 82$、$\sigma = 20$ 、$n = 100$，所以，

  $z_0 = 1.0$ 。
- 查找临界值：$\alpha = 0.05$、$Z_{0.05} = 1.645$。
- 比较并做出判断：因为 $Z_0 < Z_{0.05}$ ，因此，不能拒绝零假设，认为CJW顾客满意得分不高于80分。
- 计算p值：p=0.1587>0.05，

因此，不能拒绝零假设。

当 $\alpha = 0.05$时，$Z_{0.05}$＝?



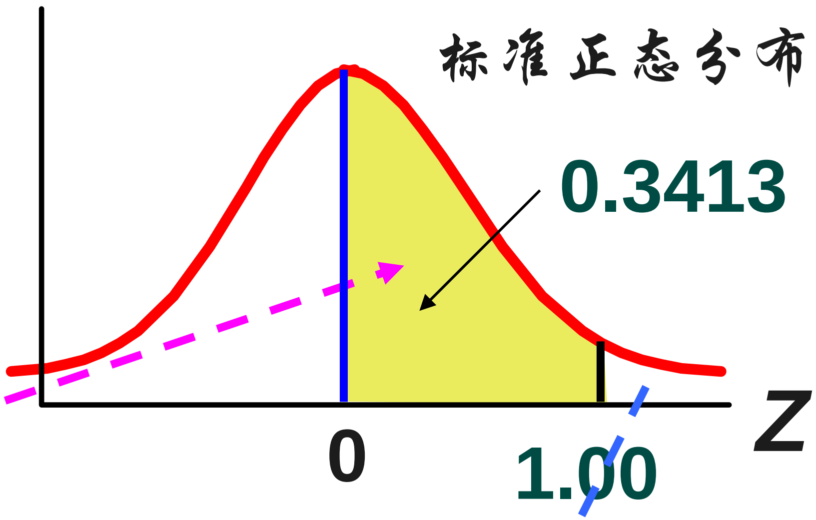
Z	.04	.05	.06
	4251	.4265	.4279
1.5	.4382	.4394	.4406
1.6	.4495	**.4505**	.4515
1.7	.4591	.4599	.4608

我会查了，$Z_{0.05}$＝1.645!

单侧检验的P值计算

标准正态分布表（部分）

Z	.00	.01	.02
0.8	.2881	.2910	.2939
0.9	.3159	.3186	.3212
1.0	**.3413**	.3438	.3461
1.1	.3643	.3665	.3686

① 

② **P=0.5－0.3413=0.1587**

4.3 假设检验（续）

□ 某公司年末审计时间(例4.1.2)的假设检验（问题2）

- 建立假设：$H_0: \mu = 18$；　$H_1: \mu \neq 18$
- 构建统计量：$t = \dfrac{\bar{x} - \mu_0}{s/\sqrt{n}} \sim t(n-1)$
- 计算检验统计量：$\bar{x} = 19.25$、$s = 5.44$、$n = 20$，所以，$t_0 = 1.0276$。

如何查表？

- 查找临界值：$\alpha = 0.05$、$t_{0.025}(19) = 2.093$。
- 比较并做出判断：因为 $|t_0| < t_{0.025}$，因此，不能拒绝零假设，认为平均时间为18小时左右。
- 计算p值：p>>0.05，
- 因此，不能拒绝零假设。

如何计算P值？

如何查t分布临界值表？

即$t_{\alpha/2}$＝？

假定：

n = 3

df= n - 1 = 2

α = .10

$\alpha/2$ =.05

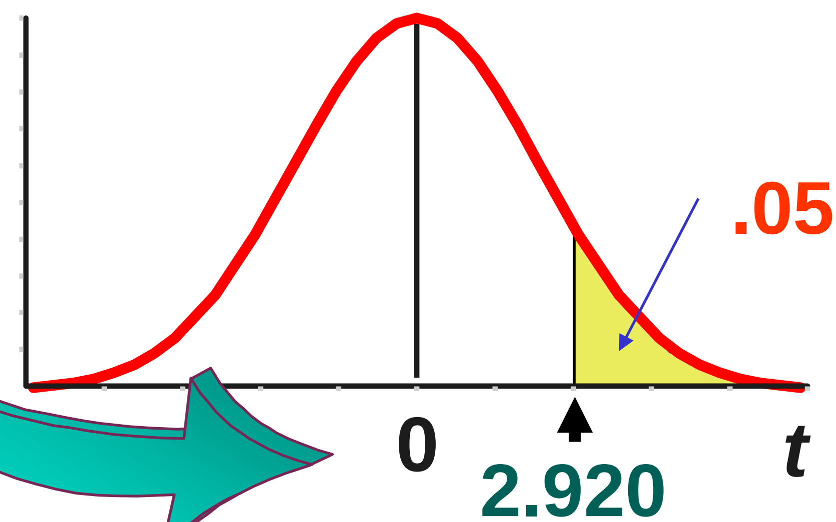
①

α / 2

②

	右侧的尾部面积		
df	.25	.10	**.05**
1	1.000	3.078	6.314
2	0.817	1.886	**2.920**
3	0.765	1.638	2.35

t的临界值



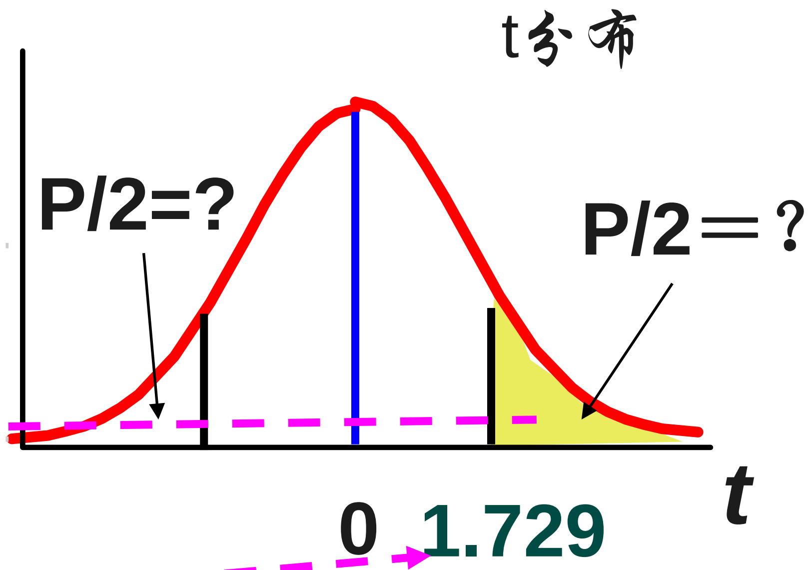
双侧检验的P值计算

t分布表（部分）

	右侧的尾部面积		
df	.05	.025	.010
18	1.734	2.101	2.552
19	1.729	2.093	2.539
20	1.725	2.086	2.528

②

①



③ P=2×0.05=0.10

4.3 假设检验（续）

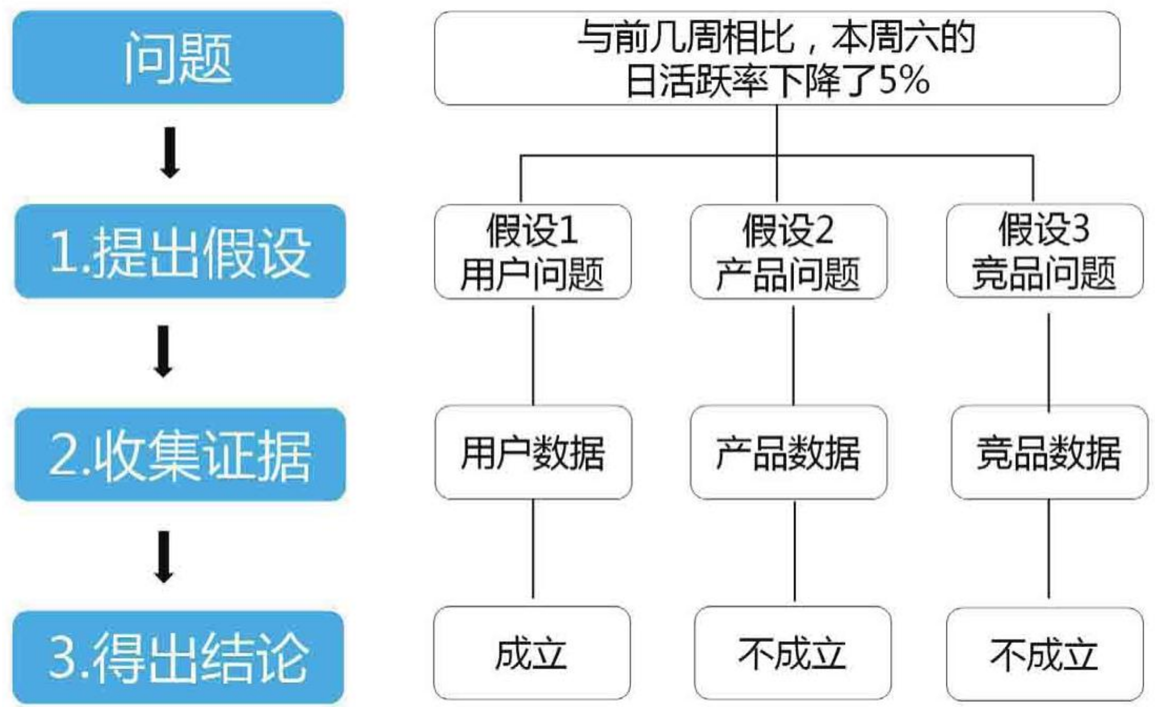
◻ 假设检验在实践中的应用

➢ 假设检验的应用范围

- 在经济、管理、法律、医学等各个领域均有广泛的应用

➢ 实践应用步骤

- 业务问题
- 提出假设
- 收集证据
- 得出结论



4.3 假设检验（续）

假设检验中的两类错误

- 第一类错误（type I error）
  - 当零假设成立时，所犯拒绝零假设的错误。犯该类错误的概率最大允许值为 α（显著性水平）。在实际中会造成严重后果，须特别注意！！
- 第二类错误（type II error）
  - 当零假设不成立时，所犯接受零假设的错误。犯错误的概率为 β。
- 两类错误的关系
  - 发生概率关系：若 α 减少，则 β 将会增加，反之亦然。两类错误概率不会同时减少，除非增加样本容量。

4.3 假设检验（续）

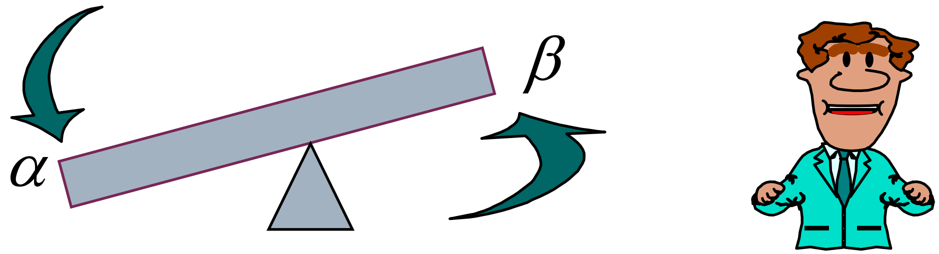


图4.3.7 两类错误概率的关系

- 两类错误与正确结论的关系

表4.3.1 两类错误与正确结论

结论	总体情况	
	H_0成立	H_0不成立
接受H_0	结论正确	第二类错误
拒绝H_0	第一类错误	结论正确

4.3 假设检验（续）

推断统计中的误差分析

- 抽样误差（变异性）
  - 随机抽样误差
  - 策略：使用大一点的样本！
- 非抽样误差（偏差）
  - 无回答误差
  - 抽样框误差
  - 计量误差
  - 策略：利用随机抽样！
- 犹如射靶一样！
  - 比较四种情况的变异性和偏差。

4.3 假设检验（续）

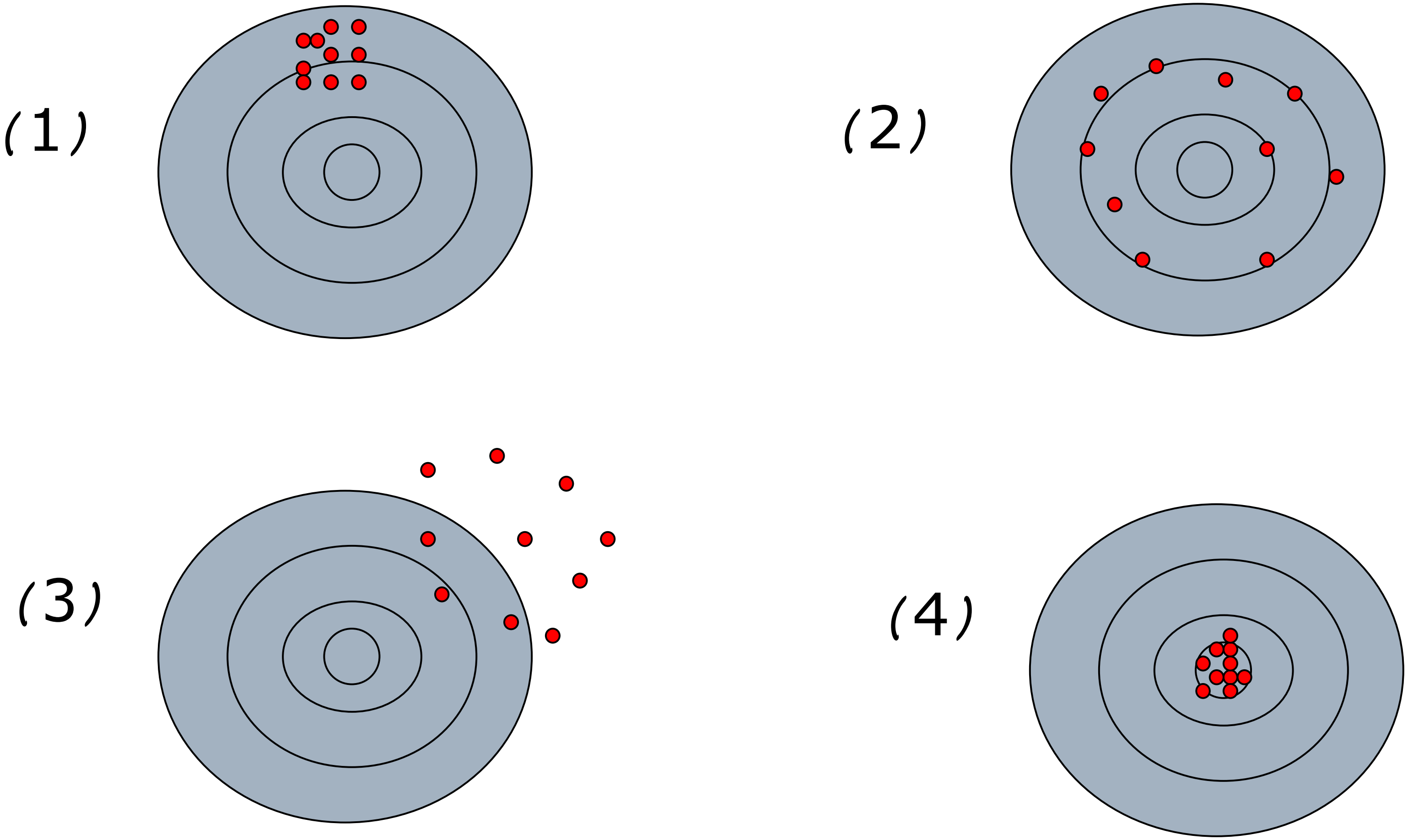


图4.3.8 射靶的偏差与变异性

简要回顾：推断统计

<table>
<tr><th>对象</th><th>目标</th><th colspan="2">条件</th><th colspan="2">形式</th><th>备注</th></tr>
<tr><td rowspan="6">样本均值</td><td>均值</td><td colspan="2">无</td><td colspan="2"></td><td rowspan="3"></td></tr>
<tr><td rowspan="2">标准差</td><td colspan="2">无限总体或n/N≤0.05的有限总体</td><td colspan="2"></td></tr>
<tr><td colspan="2">有限总体</td><td colspan="2"></td></tr>
<tr><td rowspan="3">抽样分布</td><td colspan="2"></td><td>大样本</td><td>小样本且服从正态分布</td><td rowspan="3"></td></tr>
<tr><td colspan="2">总体标准差已知</td><td></td><td></td></tr>
<tr><td colspan="2">总体标准差未知</td><td></td><td></td></tr>
<tr><td rowspan="9">总体均值</td><td>总体均值点估计</td><td>$\hat{\theta}$</td><td>总体均值的优良估计量</td><td colspan="2"></td><td></td></tr>
<tr><td rowspan="3">总体均值区间估计</td><td rowspan="3">$[\hat{\theta}-\Delta,\hat{\theta}+\Delta]$</td><td></td><td>大样本</td><td>小样本且服从正态分布</td><td rowspan="3"></td></tr>
<tr><td>总体标准差已知</td><td></td><td></td></tr>
<tr><td>总体标准差未知</td><td></td><td></td></tr>
<tr><td rowspan="5">总体均值假设检验</td><td rowspan="3">检验统计量</td><td></td><td>大样本</td><td>小样本且服从正态分布</td><td rowspan="5"></td></tr>
<tr><td>总体标准差已知</td><td></td><td></td></tr>
<tr><td>总体标准差未知</td><td></td><td></td></tr>
<tr><td rowspan="2">拒绝域、P值计算</td><td>双侧检验（H_0: $\mu_x=\mu_0$，H_1: $\mu_x\neq\mu_0$）</td><td>左侧检验（H_0: $\mu_x\geq\mu_0$，H_1: $\mu_x<\mu_0$）</td><td>右侧检验（H_0: $\mu_x\leq\mu_0$，H_1: $\mu_x>\mu_0$）</td></tr>
<tr><td></td><td></td><td></td></tr>
</table>

<table>
<tr><th>对象</th><th>目标</th><th colspan="2">条件</th><th colspan="2">形式</th><th>备注</th></tr>
<tr><td rowspan="6">样本均值</td><td>均值</td><td colspan="2">无</td><td colspan="2"></td><td rowspan="3"></td></tr>
<tr><td rowspan="2">标准差</td><td colspan="2">无限总体或n/N≤0.05的有限总体</td><td colspan="2"></td></tr>
<tr><td colspan="2">有限总体</td><td colspan="2"></td></tr>
<tr><td rowspan="3">抽样分布</td><td colspan="2"></td><td>大样本</td><td>小样本且服从正态分布</td><td rowspan="3"></td></tr>
<tr><td colspan="2">总体标准差已知</td><td></td><td></td></tr>
<tr><td colspan="2">总体标准差未知</td><td></td><td></td></tr>
<tr><td rowspan="4"></td><td>总体均值点估计</td><td>$\hat{\theta}$</td><td>总体均值的优良估计量</td><td colspan="2"></td><td></td></tr>
<tr><td rowspan="3">总体均值区间估计</td><td rowspan="3">$[\hat{\theta}-\Delta,\hat{\theta}+\Delta]$</td><td></td><td>大样本</td><td>小样本且服从正态分布</td><td rowspan="3"></td></tr>
<tr><td>总体标准差已知</td><td></td><td></td></tr>
<tr><td>总体标准差未知</td><td></td><td></td></tr>
</table>

推断统计的主要问题回顾

- 1. 点估计量的优良标准
- 2. 区间估计的定义、构成要素和基本步骤
- 3. 总体均值的置信区间形式（四种情况）
- 4. 置信区间宽度（抽样误差）的影响因素
- 5. 假设检验的定义、类型和步骤
- 6. 总体均值的假设检验统计量形式（四种情况）
- 7. 假设检验的两类错误的含义及其关系