

第三节 估计量和样本量

估计量的性质★★★★

抽样误差的估计★★

样本量计算★★★★

考点 1 估计量的性质

什么样的估计量是好估计量？

衡量的标准一般有以下三个：

- 估计量的无偏性
- 估计量的有效性
- 估计量的一致性

估计量的无偏性	【说明】 无偏性不是要求估计量与总体参数不得有偏差，既然是抽样，必然存在抽样误差，不可能与总体完全相同。 解释 无偏性指的是如果对这同一个总体反复多次抽样，则要求各个样本所得出的估计量的平均值等于总体均值。
估计量的有效性	既然是抽样，估计量与总体参数之间必然存在着一定的误差，衡量这个误差大小的一个指标就是方差，方差越小，估计量对总体的估计也就越准确，这个估计量也就越有效。
估计量的一致性	估计量的一致性，是随着样本量的增大，估计量的值如果稳定于总体参数的真值，这个估计量就有一致性，可称为一致估计量。

【单选题】在抽样估计中，随着样本量的增大，如果估计量的值稳定于总体参数的真值，这个估计量具有的性质是（ ）。

- A.一致性
- B.无偏性
- C.有效性
- D.确定性

【答案】A

【单选题】（2019）总体参数的无偏估计量的方差小于其他的无偏估计量的是（ ）。

- A.有效性
- B.一致性
- C.重要性
- D.无偏性

【答案】A

考点 2 抽样误差的估计

抽样误差虽然无法避免，但它是可计算的

抽样误差的影响因素	
影响因素	具体影响
(1) 抽样误差与总体分布有关	总体单位值之间差异越大，即总体方差越大，抽样误差越大

(2) 抽样误差与样本量 n 有关	其他条件相同，样本量越大，抽样误差越小
(3) 抽样误差与抽样方式和估计量的选择也有关	例如分层抽样的估计量方差一般小于简单随机抽样
(4) 利用有效辅助信息的估计量也可以有效的减小抽样误差	

计算公式（了解）

实践中总体方差是未知的，可以利用样本方差来估计，估计公式了解。

$$\hat{V}(\bar{y}) = \left(1 - \frac{n}{N}\right) \frac{s^2}{n}$$

(2023 补) 采取不放回简单随机抽样从总体(N=10000)抽取样本(n=100)，用样本均值 $\bar{y} = \frac{1}{100} \sum_{i=1}^{100} y_i$ ，估计总体均值 $\mu = \frac{1}{10000} \sum_{i=1}^{10000} Y_i$ ，样本方差=100，则估计量的方差为（ ）。

- A.99
- B.9.9
- C.0.99
- D.0.9

【答案】 C

【解析】 本题考查抽样抽查的估计 $V(\bar{y}) = \left(1 - \frac{n}{N}\right) \frac{S^2}{n} = \left(1 - \frac{100}{10000}\right) \times \frac{100}{100} = 0.99$ 。

【多选题】 下列关于抽样误差的表述错误的有（ ）。

- A.利用有效辅助信息的估计量也可以有效的减小抽样误差
- B.抽样误差与总体分布有关，总体方差越大，抽样误差越大
- C.其他条件相同，样本量越大，抽样误差越大
- D.抽样误差与抽样方式和估计量的选择也有关
- E.分层抽样的估计量方差一般大于简单随机抽样

【答案】 CE

【解析】 其他条件相同，样本量越大，抽样误差越小。分层抽样的估计量方差一般小于简单随机抽样。

考点 3 样本量的计算

影响因素	含义
1.调查的精度	要求的调查精度越高（误差水平越小），所需要的样本量就越大。
2.总体的离散程度	在其他条件相同情况下，总体方差越大，所需要的样本量也越大。
3.总体的规模	对于大规模的总体，总体规模对样本量的需求几乎没有影响。对小规模的总体，总体规模越大，为保证相同估计精度，样本量也要随之增大（但不是同比例的）。
4.无回答情况	无回答减少了有效样本量。在无回答率较高的调查项目中，样本量要大一些，以减少无回答带来的影响。
5.经费的制约	调查经费是影响样本量的一个十分重要的因素。事实上，样本量是调查经费与调查精度之间的某种折中和平衡。
6.其他因素	调查的限定时间，实施调查的人力资源也是影响样本量的客观因素。

【多选题】影响样本量的因素有（ ）。

- A.总体的离散程度
- B.调查的精度
- C.无回答情况
- D.经费的制约
- E.抽样方法

【答案】 ABCD

【多选题】(2019) 关于样本量的说法。正确的有（ ）。

- A.经费越少样本量越小
- B.总体规模越大样本量要同比例增大
- C.调查误差越小所需的样本量越大
- D.无回答率越高需要抽取的样本量越大
- E.总体方差越小所需的样本量越小

【答案】 ACDE

通常情况下，样本中所包含的样本单位数量（样本量，n）增大，相应样本估计量的抽样误差会减小；反之亦然。

不考虑费用限制、无回答情况及其他影响因素时，简单随机抽样的样本量计算公式为：

$$n_0 = \frac{u_\alpha^2}{d^2} S^2, n = \frac{n_0}{1 + \frac{n_0}{N}}$$

N：总体规模，当 N 很大时，对样本量 n 的影响较小，样本量 n 接近于 n_0 。

u_α ：标准正态分布的双侧 α 分位数，当置信度 $(1 - \alpha)$ 为 95% 时， u_α 值为 1.96。

d：绝对允许误差，允许误差越小，人们就越相信抽样结果接近于“真实”。

S^2 ：总体方差， S^2 未知时，一般用样本方差 S^2 替代。

在比例估计中，若 P 为总体比例，则总体方差计算公式为 $P(1-P)$

假定总体比例在 0.5 附近，总体规模为 1000 万个，当绝对允许误差分别为 0.01, 0.03, 0.05 时，简单随机抽样所需要的样本量分别为 9604、1068、385。随着样本量的增大，增加同等样本所增加的精度幅度呈下降趋势。

$$n_0 = \frac{u_\alpha^2}{d^2} S^2, n = \frac{n_0}{1 + \frac{n_0}{N}}$$

当 $d=0.01$, $u_\alpha=1.96$

$S^2=p(1-p)=0.5 \times (1-0.5)=0.25$

$N_0 = \frac{1.96^2}{0.01^2} \times 0.25 = \frac{3.8416}{0.0001} \times 0.25 = 9604$

(2023 补) 在置信度 $(1-\alpha)$ 下用随机样本以估算量估计总体参数 θ ，绝对允许误差 d 的含义是（ ）。

- A. $\Pr(|\hat{\theta} - \theta| \leq d) = 1 - \alpha$

B. $\Pr(|\hat{\theta} - \theta| \geq d) = 1 - \alpha$

C. $\Pr(|\hat{\theta} - \theta| \geq 1 - \alpha) \geq d$

D. $\Pr(|\hat{\theta} - \theta| \leq d) \geq 1 - \alpha$

【答案】 A

【解析】 本题考查样本量的计算。在样本量计算中， d 为绝对允许误差，用于表示一定的置信度 $(1-\alpha)$ 下样本估计值与总体参数 θ 之间的误差不超过某一给定的最大可能范围，允许误差越小，表明人们就越相信抽样结果接近于“真实”。