

第三节 估计量和样本量

(一) 估计量的性质★★★★

(二) 抽样误差的估计★★

(三) 样本量的计算★

(一) 估计量的性质★★★★

性质	内涵
无偏性	对于不放回简单随机抽样，所有可能的样本均值取值的平均值总是等于总体均值，这就是样本均值估计量的无偏性。
有效性	在同一抽样方案下，对某一总体参数 θ ，如果有两个无偏估计量 $\hat{\theta}_1$ 和 $\hat{\theta}_2$ ，由于样本的随机性， $\hat{\theta}_1$ 的可能样本取值较更密集在总体参数真值 θ 附近，人们会认为 $\hat{\theta}_1$ 比 $\hat{\theta}_2$ 更有效， $\hat{\theta}_1$ 的方差小于 $\hat{\theta}_2$ 的方差。
一致性	随着样本量的增大，估计量的值如果稳定于（或收敛于）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。

样本	样本均值
6, 7	6.5
6, 8	7
6, 9	7.5
6, 10	8
7, 8	7.5
7, 9	8
7, 10	8.5
8, 9	8.5
8, 10	9
9, 10	9.5

【2023 年·单选】对于不放回简单随机抽样，所有可能的样本均值取值的平均值总是等于总体均值，这就是样本估计量的（ ）。

A. 无偏性

- B. 有效性
- C. 一致性
- D. 渐进性

答案：A

解析：无偏性指所有可能的样本均值取值的平均值总是等于总体均值。

【2024 年 · 单选】在不放回简单随机抽样中，估计量的无偏性是指（ ）。

- A. 所有可能样本估计值的平均值等于总体参数
- B. 所有可能样本估计值都等于总体参数
- C. 所有可能样本估计值集中在总体参数附近
- D. 随着样本量的增大，样本估计值趋向总体参数

答案：A

解析：无偏性指所有可能样本估计值的平均值等于总体参数，故 A 正确。

【2024 年 · 单选】假设某总体参数 θ 有多个无偏估计量 $\hat{\theta}_1$ 和 $\hat{\theta}_2$ ，已知 $\text{Var}(\hat{\theta}_1) > \text{Var}(\hat{\theta}_2)$ ，这 2 个估计量的关系是（ ）。

- A. $\hat{\theta}_1$ 比 $\hat{\theta}_2$ 更一致
- B. $\hat{\theta}_2$ 比 $\hat{\theta}_1$ 更一致
- C. $\hat{\theta}_2$ 比 $\hat{\theta}_1$ 更有效
- D. $\hat{\theta}_1$ 比 $\hat{\theta}_2$ 更有效

答案：C

解析：在同一抽样方案下，对某一总体参数 θ ，如果有两个无偏估计量 $\hat{\theta}_1$ 和 $\hat{\theta}_2$ ，由于样本的随机性， $\hat{\theta}_2$ 的可能样本取值较更密集在总体参数真值 θ 附近，人们会认为 $\hat{\theta}_2$ 比 $\hat{\theta}_1$ 更有效， $\hat{\theta}_2$ 的方差小于 $\hat{\theta}_1$ 的方差。

【2023 年 · 单选】估计量的一致性是指（ ）。

- A. 随着样本量增大，估计量的值稳定于总体参数真值
- B. 估计值等于总体参数真值
- C. 所有可能样本的估计量均值等于总体参数真值
- D. 估计量的方差最小

答案：A

解析：一致性是指随着样本量的增大，估计量的值如果稳定于（或收敛于）总体参数的真值。C 是无偏性，D 是有效性。

(二) 抽样误差的估计★★

估计量的方差的计算	估计量的方差 $V(\bar{y}) = (1 - \frac{\text{样本量}n}{\text{总体个数}N}) \times \frac{\text{总体方差}S^2}{\text{样本量}n}$
抽样误差的来源	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，抽样误差就越大。 ②抽样误差与样本量n有关，在其他条件相同情况下，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系。 ④利用有效辅助信息也可以有效地减小抽样误差。

从某个 $N=1000$ 的总体中抽出一个样本容量为100的不放回简单随机抽样，样本均值 $\bar{y}=50$ ，总体方差 $S^2=200$ ，则对总体均值的估计 $\hat{Y}=\bar{y}=50$ ，估计量 $\bar{y}$ 方差的估计为：

$$\text{估计量的方差} V(\bar{y}) = (1 - \frac{\text{样本量}n}{\text{总体个数}N}) \times \frac{\text{总体方差}S^2}{\text{样本量}n}$$

$$V(\bar{y}) = \left(1 - \frac{100}{1000}\right) \times \frac{200}{100} = 1.8$$

【例题·单选】某总体的总体规模 $N=10000$ ，总体方差 $S^2=1000$ ，采取不放回简单随机抽样方法抽取样本容量为100，则估计量方差 $V_{100}(\bar{y})$ 是（ ）。

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

答案：A

解析： $V_{100}(\bar{y}) = (1 - 100/10000) \times 1000/100 = 9.9$ 。

【2021年·单选】某总体的总体规模 $N=1000000$ ，总体方差 $S^2=1000$ ，采取不放回简单随机抽样方法分别抽取样本容量为100和样本容量为1000的两个样本，并各自用样本均值 $\bar{y}$ 估计总体均值 $\bar{Y}$ ，这两个样本的均值估计量方差 $V_{100}(\bar{y})$ 和 $V_{1000}(\bar{y})$ 的关系是（ ）。

- A. $V_{100}(\bar{y}) \approx 10V_{1000}(\bar{y})$
- B. $10 V_{100}(\bar{y}) \approx V_{1000}(\bar{y})$
- C. $100V_{100}(\bar{y}) \approx V_{1000}(\bar{y})$
- D. $V_{100}(\bar{y}) \approx 100V_{1000}(\bar{y})$

答案：A

解析： $V_{100}(\bar{y}) = (1-100/1000000) \times 1000/100$
； $V_{1000}(\bar{y}) = (1-1000/1000000) \times 1000/1000$ 。经计算可得： $V_{100}(\bar{y}) \approx 10V_{1000}(\bar{y})$ 。

（三）样本量的计算★

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
在实际应用中，样本量的确定主要应考虑以下几方面因素。

1. 调查的精度（精度越高，样本量就越大）。
2. 总体的离散程度（总体方差越大，样本量越大）。
3. 总体的规模

小规模总体，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（无回答率较高，样本量大）。
5. 经费的制约。

简单随机抽样的样本量计算公式

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

【例题·单选】某市开展居民垃圾分类知晓率调查，采用简单随机抽样。预计总体比例 P 约为0.5，置信度为95%（ $u=1.96$ ），总体规模极大（ $N>10000$ ）。若要求绝对允许误差 d 不超过3%，则至少需要的样本量

最接近（ ）。

A. 385

B. 1068

C. 2401

D. 9604

答案：B

解析： $S^2 = P(1 - P) = 0.5 \times 0.5 = 0.25, d =$

0.03。代入公式：

$$n \approx n_0 = \frac{u_{\alpha}^2}{d^2} S^2 = \frac{1.96^2 \times 0.25}{0.03^2} = \frac{0.9604}{0.0009} \approx 1067.1$$

框架导航

