



浙江财经大学
Zhejiang University of Finance & Economics

区间估计



教师: 李金昌
浙江财经大学 数据科学学院



一、点估计

点估计就是以某个样本观测数据为依据，对总体参数做出确定的估计，也就是用一个样本统计值来估计总体参数值。

优点：可以给出明确的数值；

缺点：正确的把握程度较低。



二、区间估计

区间估计是用一个具有一定可靠程度的**区间范围**来估计总体参数。

区间估计的两个基本要求是置信度和精确度。

区间越大，覆盖总体参数值的可能性越高，但其精确度越低。

因此区间估计要权衡置信度和精确度之间关系。

$$\pi \left(\hat{\theta} - \underbrace{\Delta}_{\text{精确度}} < \theta < \hat{\theta} + \underbrace{\Delta}_{\text{置信度}} \right) = 1 - \alpha$$



区间估计的步骤

- (1) 根据样本数据计算 $se(\bar{x})$;
- (2) 给定置信水平 $1 - \alpha$, 查表得到 $Z_{\alpha/2}$
- (3) 总体均值的估计区间 $\Delta = z_{\alpha/2} se(\bar{x})$

重复抽样：
$$\left[\bar{x} - Z_{\alpha/2} \cdot \sqrt{\frac{S^2}{n}}, \bar{x} + Z_{\alpha/2} \cdot \sqrt{\frac{S^2}{n}} \right]$$

不重复抽样：
$$\left[\bar{x} - Z_{\alpha/2} \cdot \sqrt{\frac{S^2}{n} \left(1 - \frac{n}{N}\right)}, \bar{x} + Z_{\alpha/2} \cdot \sqrt{\frac{S^2}{n} \left(1 - \frac{n}{N}\right)} \right]$$



举例

某企业生产某种产品的工人有1000人，某日采用不重复抽样从中随机抽取100人调查他们的当日产量。要求在95%概率保证程度下，估计该厂全部工人的人均日产量。



区间估计

按日产量分组 (件)	工人数 (人)
110~114	3
114~118	7
118~122	18
122~126	23
126~130	21
130~134	18
134~138	6
138~142	4
合计	100



第一步，计算抽样标准误差

$$\bar{x} = \frac{\sum x_i f_i}{\sum f_i} = \frac{12600}{100} = 126 \text{ (件)}$$

$$s = \sqrt{\frac{\sum (x_i - \bar{x})^2 f_i}{\sum f_i - 1}} = \sqrt{\frac{4144}{99}} = 6.47 \text{ (件)}$$

$$\begin{aligned} se(\bar{x}) &= \sqrt{\frac{s^2}{n}(1-f)} \\ &= \sqrt{\frac{6.47^2}{100} \left(1 - \frac{100}{1000}\right)} = 0.614 \text{ (件)} \end{aligned}$$



第二步，查表计算抽样极限误差

$$\because 1-\alpha = 95\%$$

$$\therefore Z_{\alpha/2} = 1.96$$

$$\therefore \Delta = Z_{\alpha/2} \cdot se(\bar{x}) = 1.96 \times 0.614 = 1.203 (\text{件})$$



区间估计

第三步，根据点估计和抽样极限误差，得到区间估计

$$\bar{X} \in [\bar{x} - \Delta, \bar{x} + \Delta]$$

$$\Rightarrow 126 - 1.203 \leq \bar{X} \leq 126 + 1.203$$

$$\Rightarrow [124.80, 127.20]$$

所以全部工人的人均日产量处于124.80和127.20件之间。



三、样本容量的确定

通常，在抽样调查之间，需要确定样本容量。样本容量大小受总体分布、对抽样精度和可靠程度的要求、抽样方法和调查经费等因素的影响。



在不考虑调查经费时，简单随机抽样的样本容量可由下式计算：

$$n = \frac{Z_{\alpha/2}^2 \times S^2}{\Delta^2}$$

其中总体方差 S^2 是未知的，需要采用实验研究、比较研究、全距法推断得到，通常用样本方差代替。



举例

某市拟对全市居民家庭生活收支情况进行调查，要求保证95%的可靠程度，允许误差不超过0.2万元，据去年的调查，户年收入方差为1.6万元，至少应抽多少户？



区间估计

解：根据题意，可知抽样极限误差 $\Delta = 0.2$,

总体方差 $S^2 = 1.6$

$$1 - \alpha = 95\% \rightarrow Z_{\alpha/2} = 1.96$$

$$n = \frac{Z_{\alpha/2}^2 S^2}{\Delta^2} = \frac{1.96^2 \times 1.6}{0.2^2} = 153.66 \approx 154 (\text{户})$$

即，至少应抽154户家庭。