

# 数理统计

ImageNature

May 11, 2023

抛开实际背景，总体就是一堆有大有小的数，因此用一个概率分布去描述和归纳总体是恰当的. 某种意义上来说，总体就是一个分布，它的数量指标就是服从分布的随机变量. 因此从总体中抽样和从某分布中抽样是同一个意思.

## 1 绪论

1. 若将样本观测值有小到大进行排列，记作  $X_{(1)}, \dots, X_{(n)}$ ，则称为有序样本，且可以用有序样本定义经验分布函数. 有序样本对应的是次序统计量.
2. 统计量是一类函数，统计量的分布称为抽样分布. 尽管统计量不依赖于未知参数，但是它的分布是依赖于未知参数的.
3. 设  $x_1, \dots, x_n$  是来自某个总体的样本， $\bar{x}$  是样本均值

- 若总体分布是  $N(\mu, \sigma^2)$ , 则  $\bar{x}$  的精确分布为  $N(\mu, \sigma^2/n)$
- 若总体分布不是正态分布或未知,  $E(x) = \mu, \text{Var}(x) = \sigma^2$  存在, 则当  $n$  较大时  $\bar{x}$  的极限分布 (渐进分布) 为  $N(\mu, \sigma^2/n)$ . 这里渐进分布是  $n$  较大时的近似分布.

4. 样本方差时度量样本散布大小的统计量, 样本方差定义为

$$s_n^2 = \frac{1}{n} \sum_{i=1}^n (x_i - \bar{x})^2$$

为了方便地构造无偏统计量, 一般会定义为

$$s^2 = \frac{1}{n-1} \sum_{i=1}^n (x_i - \bar{x})^2$$

在定义中,  $n$  为样本量,  $n-1$  称为偏差平方和的自由度. 含义是: 在  $\bar{x}$  确定后,  $n$  个偏差  $x_1 - \bar{x}, \dots, x_n - \bar{x}$  只有  $n-1$  个偏差可以自由变动, 因为其和为 0.