

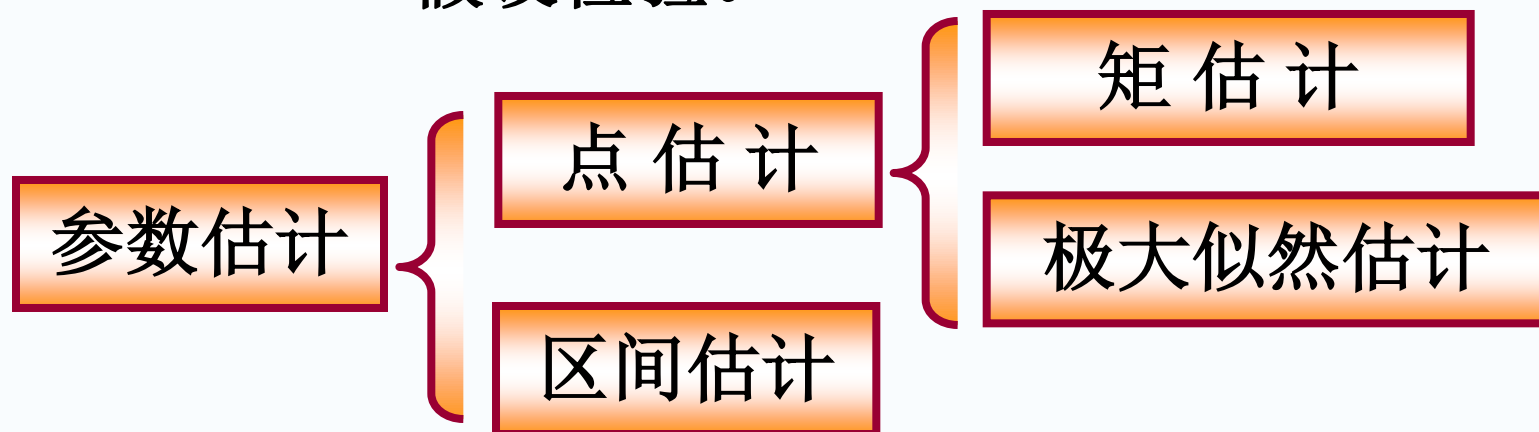


统计推断三个方面：

抽样分布(精确分布)；

参数估计；(已知分布类型)

假设检验。





§ 7.3 区间估计

由于样本的随机性, 点估计有以下缺陷:

- (1) 无从断定估计值是否为待估参数的真实值
(即使估计量是无偏有效估计量).
- (2) 不能把握估计值与参数真实值的偏离程度及估计的可靠程度.

改进:

随机区间

对于 θ 的估计, 给定一个范围 $[\hat{\theta}_1, \hat{\theta}_2]$ 满足:

- (1) $P\{\theta_1 \leq \theta \leq \theta_2\}$ 应尽可能大, 即可靠程度高;
- (2) $\hat{\theta}_2 - \hat{\theta}_1$ 应尽可能小, 精确度高。





第七章 参数估计——区间估计

定义：（双侧）

设总体的未知参数为 θ ，由样本 X_1, \dots, X_n 确定两个

统计量 $\hat{\theta}_1 = \hat{\theta}_1(X_1, \dots, X_n)$ 和 $\hat{\theta}_2 = \hat{\theta}_2(X_1, \dots, X_n)$,

对于给定的实数 $\alpha(0 < \alpha < 1)$ ，满足

$$P\{\hat{\theta}_1(X_1, \dots, X_n) \leq \theta \leq \hat{\theta}_2(X_1, \dots, X_n)\} = 1 - \alpha$$

则称随机区间 $[\hat{\theta}_1, \hat{\theta}_2]$ 为 θ 的置信度为 $1-\alpha$ 的置信区间。

$1-\alpha$ 又称置信系数或置信概率或置信水平

α 又称显著性水平，通常取值为0.01, 0.05

注：反复抽多次，每个样本值确定一个区间，这些区间中要么包含待估参数的真值，要么不包含待估参数的真值，按伯努利大数定律，其中包含待估参数的真值占 $(1-\alpha)\%$ ，不包含待估参数真值仅占 $\alpha\%$ 。





思考：应如何理解概率式 **估计**——区间估计

$$P\{\hat{\theta}_1(X_1, \dots, X_n) \leq \theta \leq \hat{\theta}_2(X_1, \dots, X_n)\} = 1 - \alpha$$

注：（1）上式应理解为随机区间 $[\hat{\theta}_1, \hat{\theta}_2]$ 以 $1 - \alpha$ 的概率包含着待估参数 θ . (**区间估计本质所在**)

$1 - \alpha$ 反映了区间估计的**可靠程度**.

（2）随机区间 $[\hat{\theta}_1, \hat{\theta}_2]$ 的长度 $\hat{\theta}_2 - \hat{\theta}_1$ 是随机变量，它反映了区间估计的**精确程度**.

希望精度与可靠程度均高，但这二者是**矛盾**的。

实际应用中广泛接受的原则是：(统计学家奈曼Neyman)

在置信度达到一定要求前提下，

寻求精确度尽可能高的区间估计，

也就是寻求区间平均长度尽可能短。

置信区间的**枢轴变量法**：正态分布中 μ 的区间估计





第七章 参数估计—区间估计

本次课的主要内容：第七章 参数估计-优良估计量及区间估计



文森特·威廉·梵高

《盛开的杏花》(1890.春)

荷兰阿姆斯特丹梵高美术馆

1890年，正是人生最潦倒的最后一年，梵高住到了阿尔勒的精神疗养院中，除了弟弟提奥和弟媳乔安娜外，无人问津。

阿尔勒色彩明亮、天空湛蓝。这个春天繁花满树，对梵高来说，意味着新生和希望。令梵高最快乐的事，恰恰也是这一年，弟弟提奥唯一的孩子的出生。这个孩子被命名为：文森特·威廉·梵高。

欣喜不已的梵高，虽然觉得自己一事无成，在给提奥的信中建议“提奥”是一个更好的名字，并马上为自己亲爱的侄子画了这一副《杏花》。

这幅画作也成为了梵高家族“唯一不肯出售的画作”。



下次课内容：第七章 参数估计-区间估计及第八章假设检验开头

