

3-5-17 较好的参数估计量

我们分析数据的目的是估计总体分布，进而推测未知的数据。

假设总体服从正态分布，那么只要估计出正态分布的参数就能得知总体分布了，所以参数估计是核心。

正态分布的参数有 2 个，分别为均值和方差。

一般来说，正态分布的参数难以直接得到，所以我们用样本均值和无偏方差作为其参数。

从本节的各个程序模拟可知，样本均值和无偏方差具有适合作为参数估计量的性质。

样本均值的均值和总体均值基本相等，样本的无偏方差的均值和总体方差基本相等。也就是说，样本均值和无偏方差都具有无偏性。

另外，样本容量越大，样本均值越接近总体均值，无偏方差也越接近总体方差。也就是说，样本均值和无偏方差具有一致性^①。

3-5-18 补充 大数定律

所谓大数定律，就是样本容量越大，样本均值越接近总体均值。

大数定律分为弱大数定律和强大数定律，具体请参考相关文献。

3-5-19 补充 中心极限定理

对于任意总体分布，样本容量越大，随机变量的和的分布越接近正态分布，这就是中心极限定理。

^① 这个性质是基于样本容量无穷大的样本得到的，程序模拟的结果只是说明它们有这种趋势，严密理论请参考本书末尾的参考文献。