1. 中心极限定理

中心极限定理是概率论中最著名的结果之一。它提出，大量的独立随机变量之和具有近似于正态的分布。因此，它不仅提供了计算独立随机变量之和的近似概率的简单方法，而且有助于解释为什么有很多自然群体的经验频率呈现出钟形(即正态)曲线这一事实。

样本均值的抽样分布是所有样本均值形成的分布，即μ的概率分布。根据中心极限定理，μ的概率分布是均值为总体均值、方差为总体方差除以样本容量的正态分布。

设样本容量为n，当n很小时，中心极限定理不能很好的发挥作用。经验表明，当n大于等于30时中心极限定理才能很好的发挥作用。n越大，抽样分布的方差越小，样本越集中。

均值标准误差也就是样本均值抽样分布的标准误差，它等于总体标准差除以根号n，n为样本容量。

1. 置信区间和置信度

一般我们用中括号[a,b]表示样本估计总体平均值误差范围的区间。a、b的具体数值取决于你对于“该区间包含总体均值”这一结果的可信程度，因此[a,b]被称为置信区间。

一般来说，选定某一个置信区间，我们的目的是为了让“ab之间包含总体平均值”的结果有一特定的概率，这个概率就是所谓的置信水平。

例如我们最常用的95%置信水平，就是说做100次抽样，有95次的置信区间包含了总体均值。

1. 伯努利分布

伯努利分布是一个离散概率分布，又名两点分布或0-1分布。伯努利试验是只有两种可能结果的单次随机试验。设伯努利试验成功的概率为p，失败的概率为q=1-p。则伯努利分布的均值为p，方差为p(1-p)。

1. t分布

当样本容量很小时（一般认为小于30时），样本均值抽样分布不应该采用正态分布，而应该采用t分布，可以认为t分布式为小样本容量置信区间的更好估计所设计的。