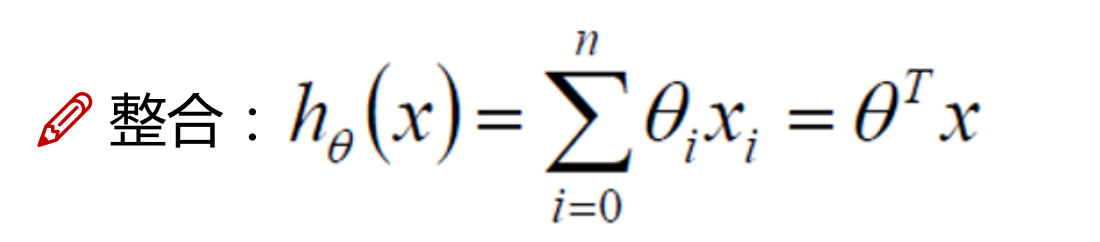
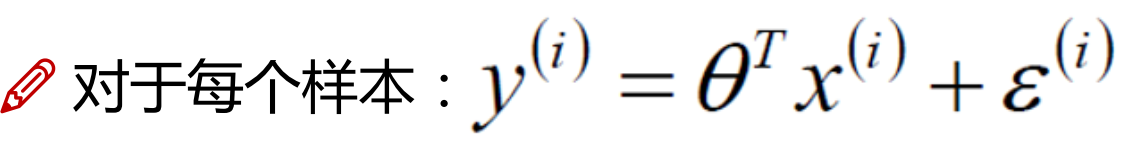
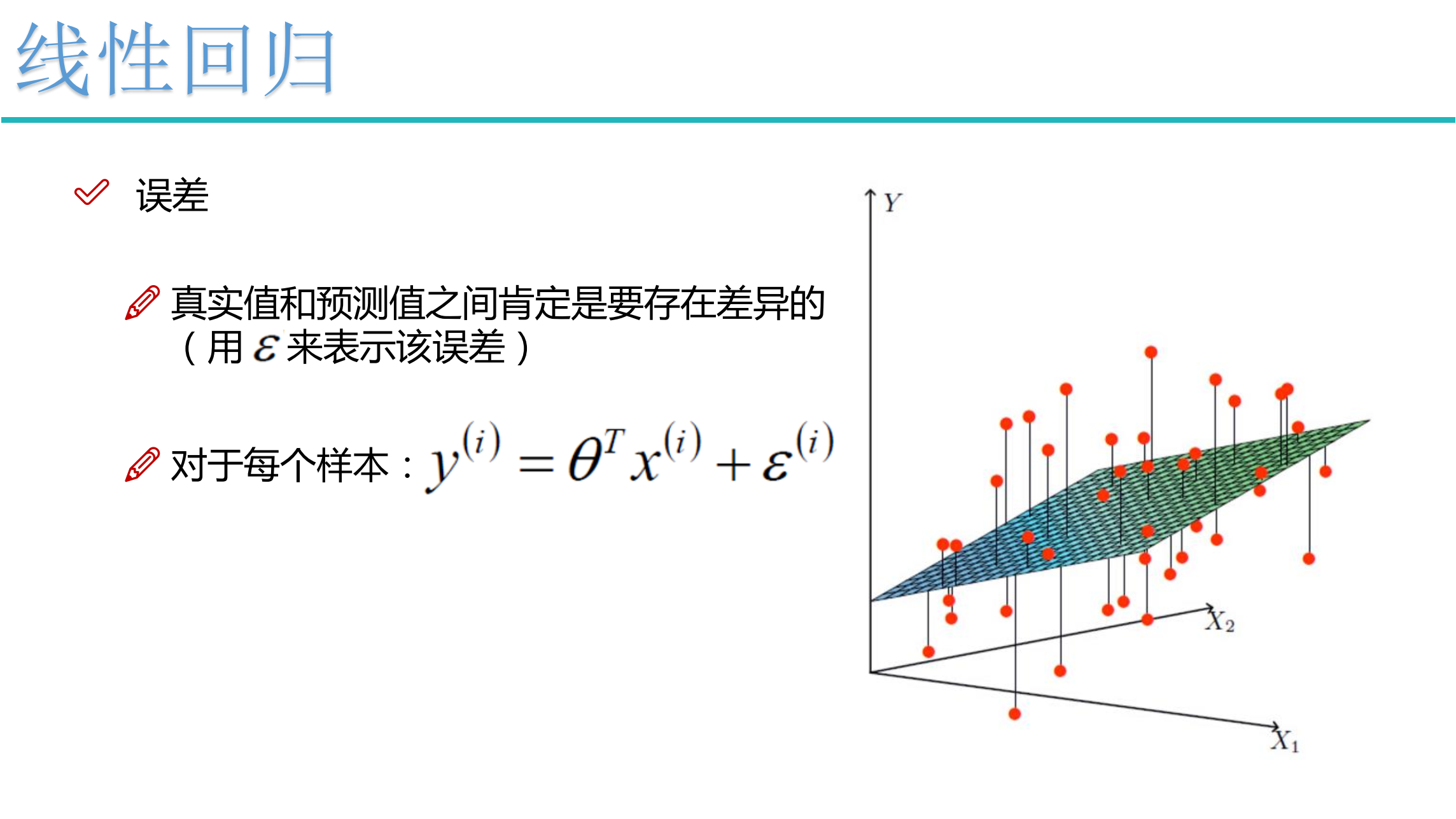


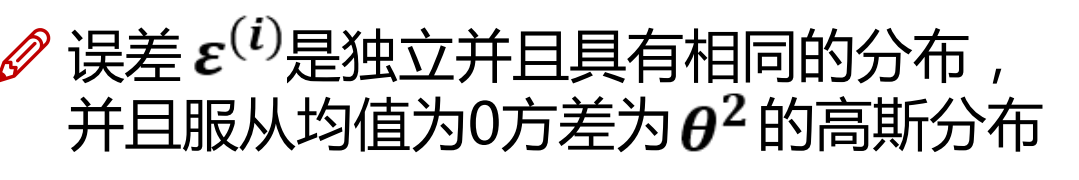
某一样本的预测结果可以表示为

对于所有样本我们可以用，这里的i和上面的i不同，

上面的i是维度，这里的i是代表样本

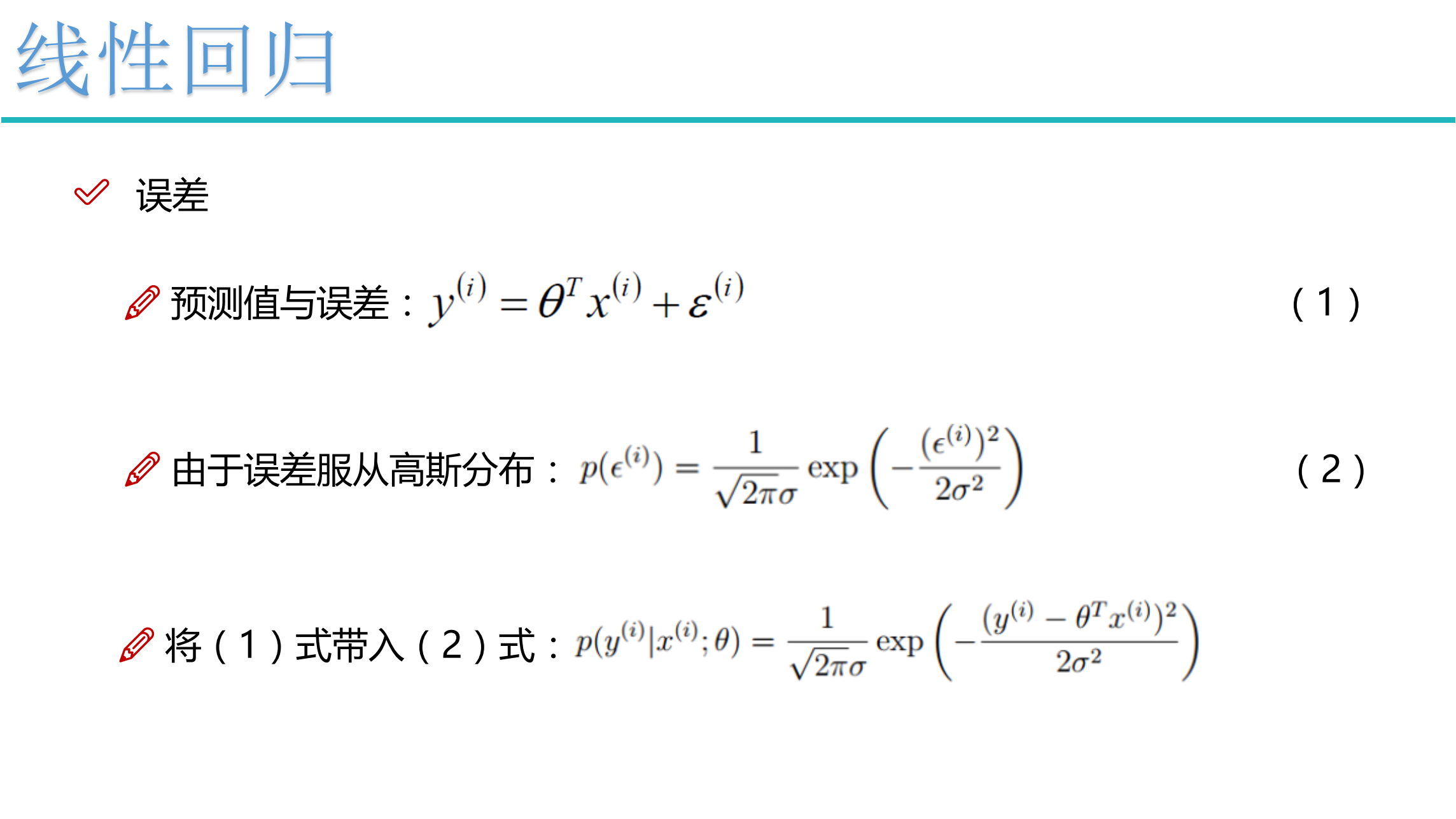






因为我们已经知道这种误差是服从某一分布的，我们就可以用似然估计来预估参数，这个参数就是分布的方差啊期望啊，

我们知道误差服从高斯分布，但是我们不知道具体的分布参数，这个需要估计

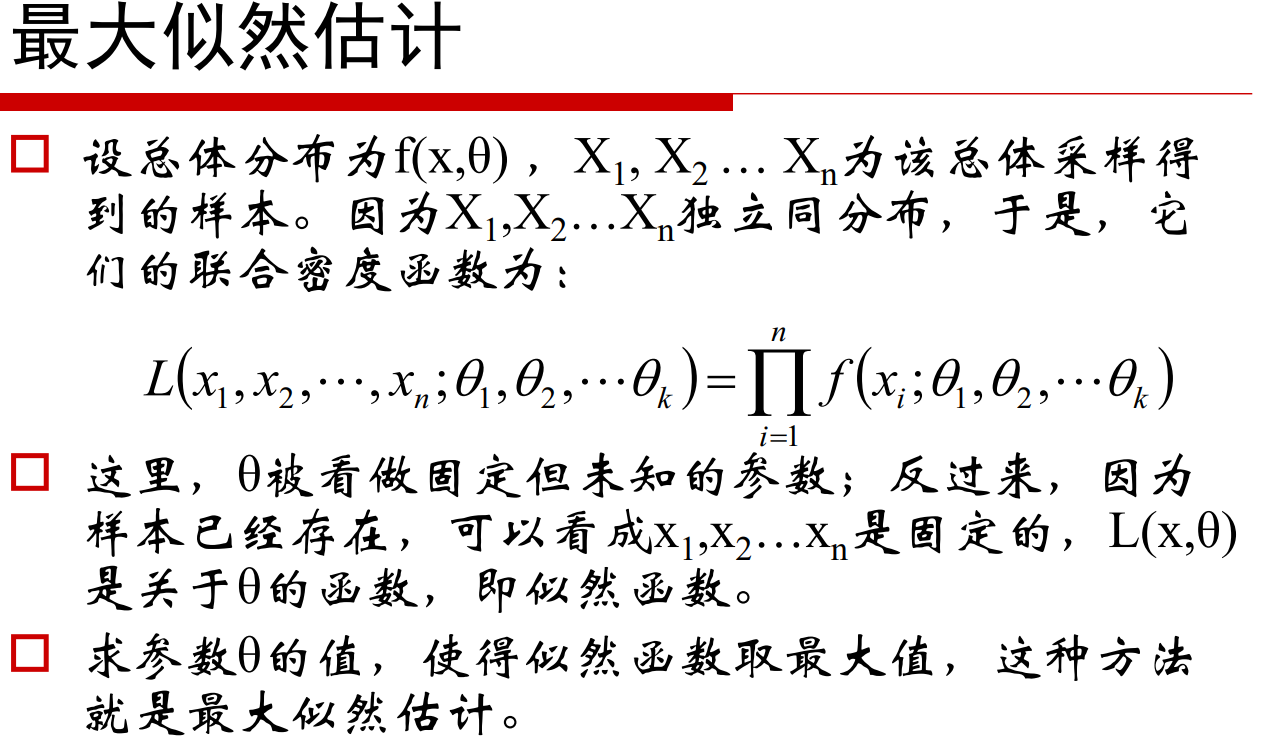


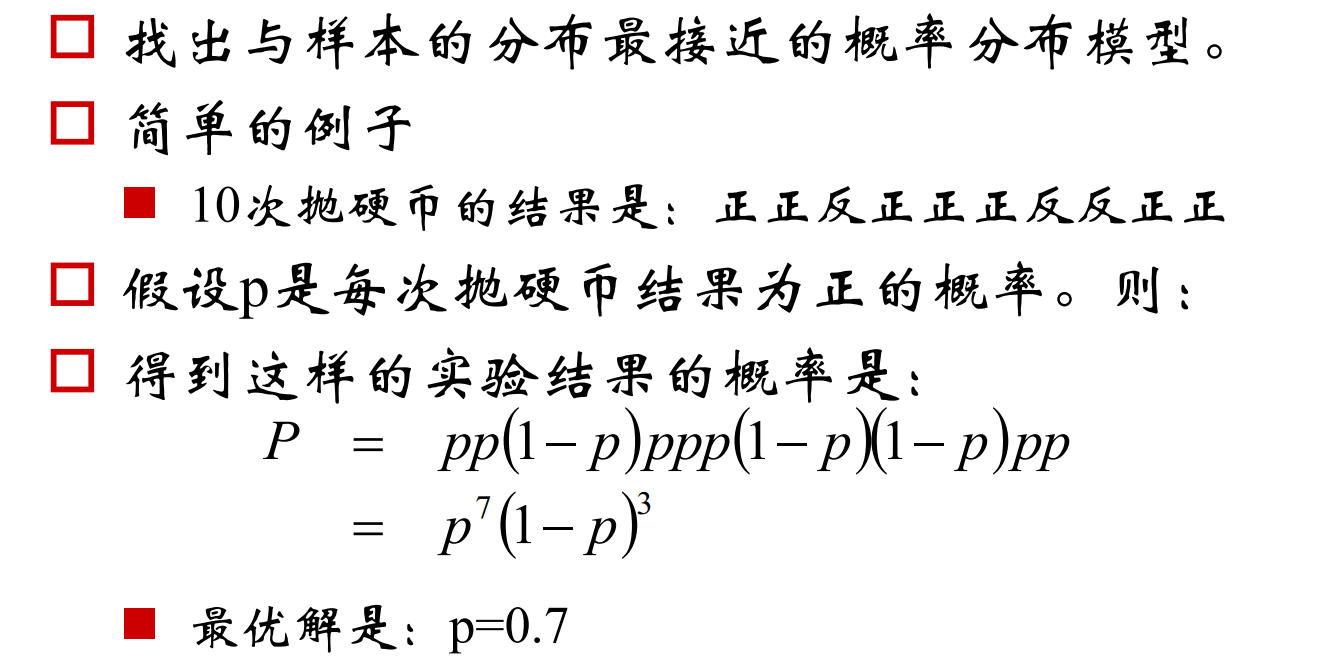
为什么用最大似然估计呢？

当我们有一个样本分布数据，我们是可以去估计某一个事件发生的概率。

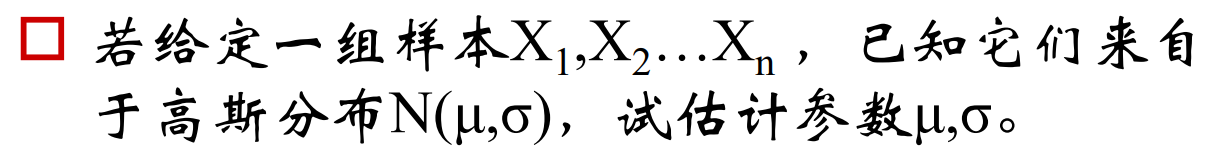
最大似然估计的通俗理解就是使当前样本最有可能发生的情况下，这个分布的参数是啥，

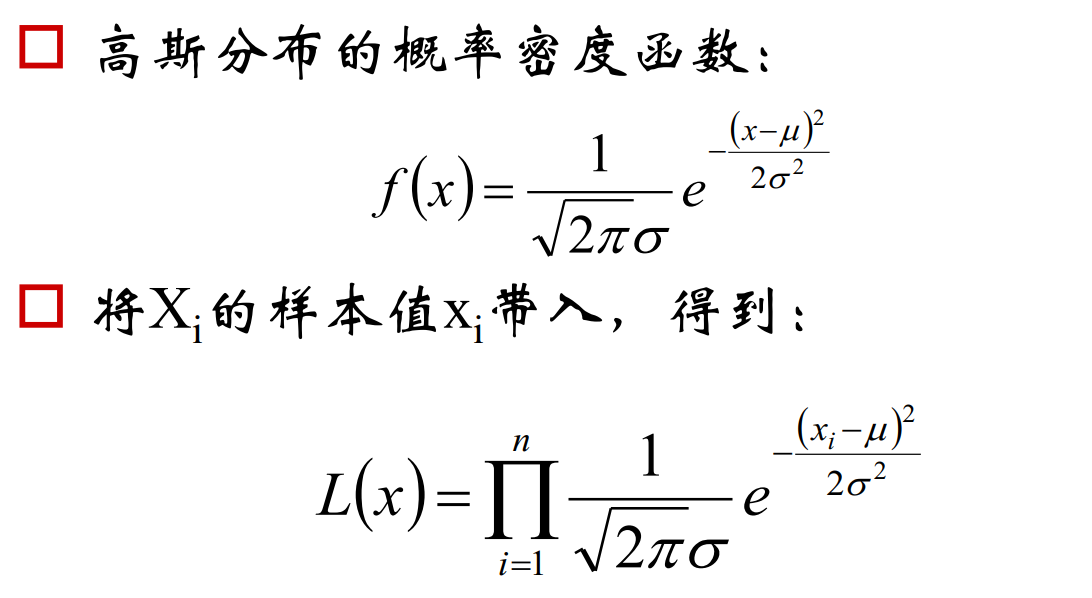
这里的参数并不是说一个样本有一个参数，而是说这个分布可能有多个参数，比如说高斯分布



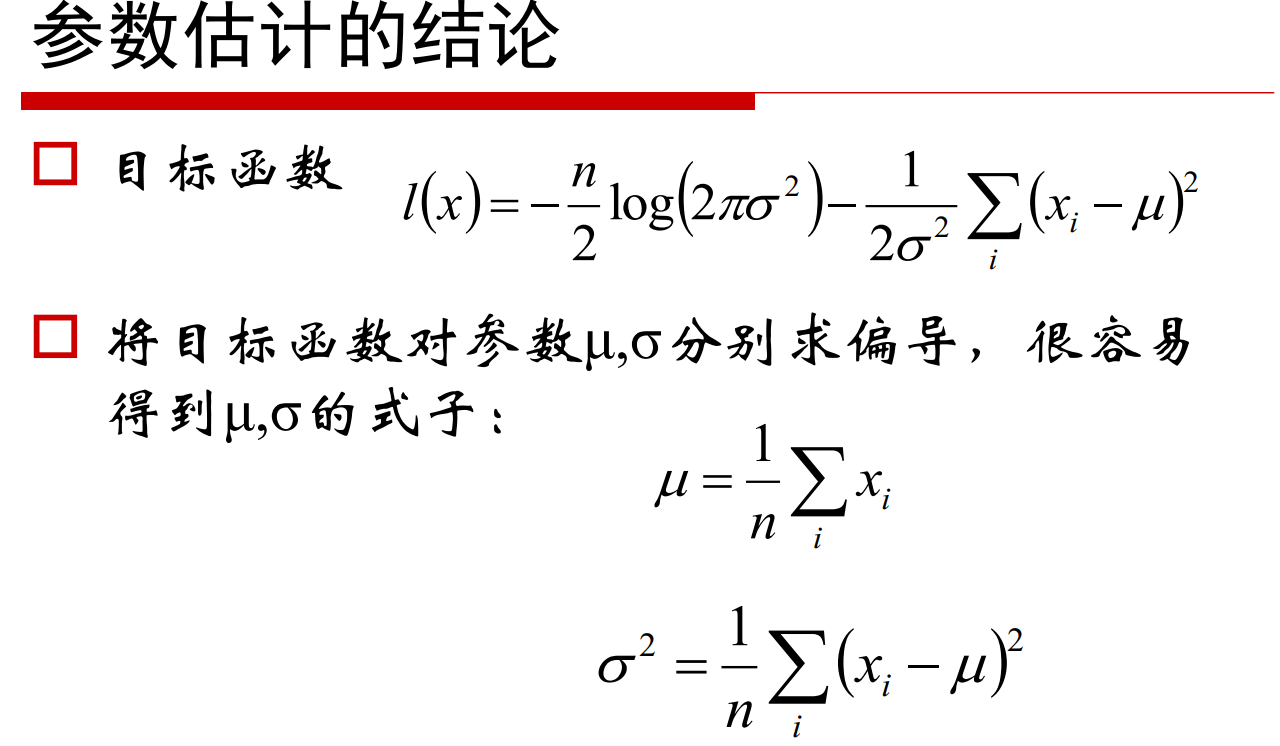


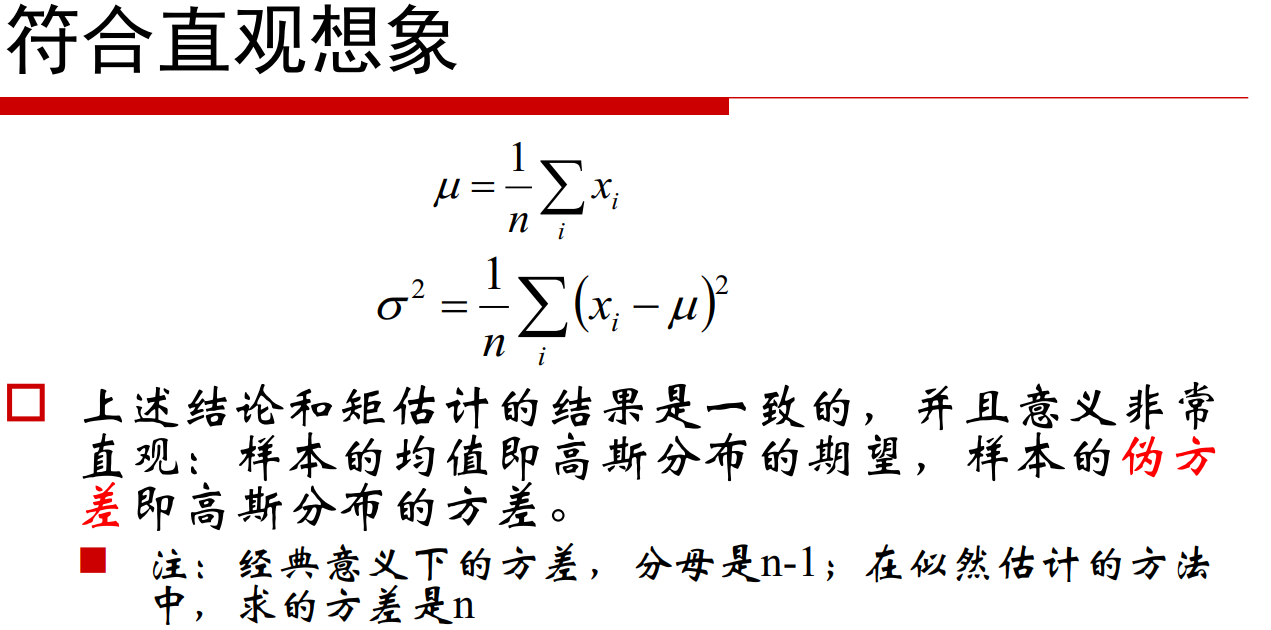
对于上例，抛硬币来说，我们得到了一个分布，虽然我们不知道这个结果服从什么已知的分布，但是我们可以用联合密度函数来预估某一事件发生的概率，也就是说p=0.7时最有可能是上述样本情况发生





F（x）代表x发生的概率，L（x）代表这些独立同分布样本连续发生的概率，最终可以用一个函数表示，这里自变量已经变成了这些参数



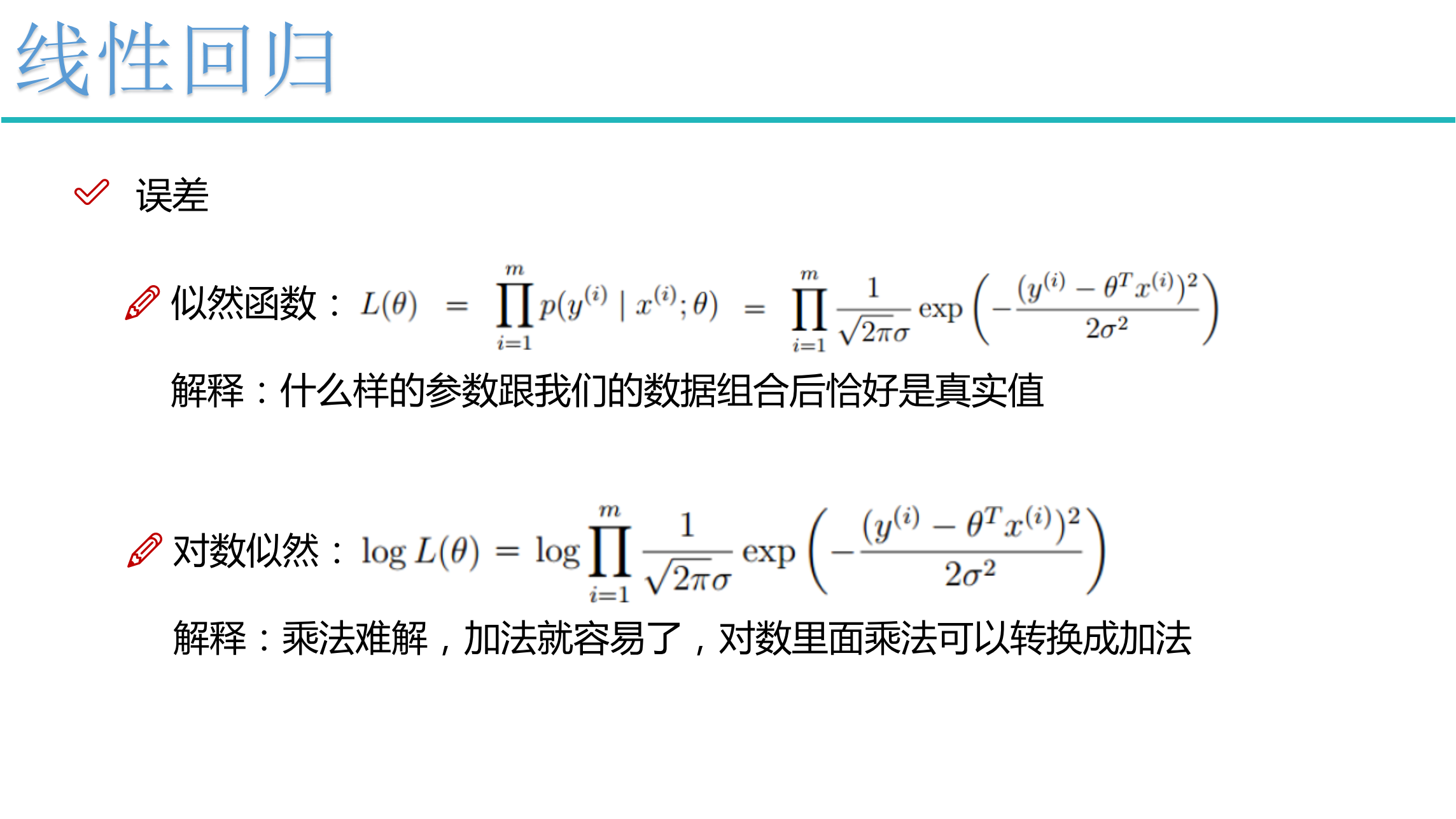


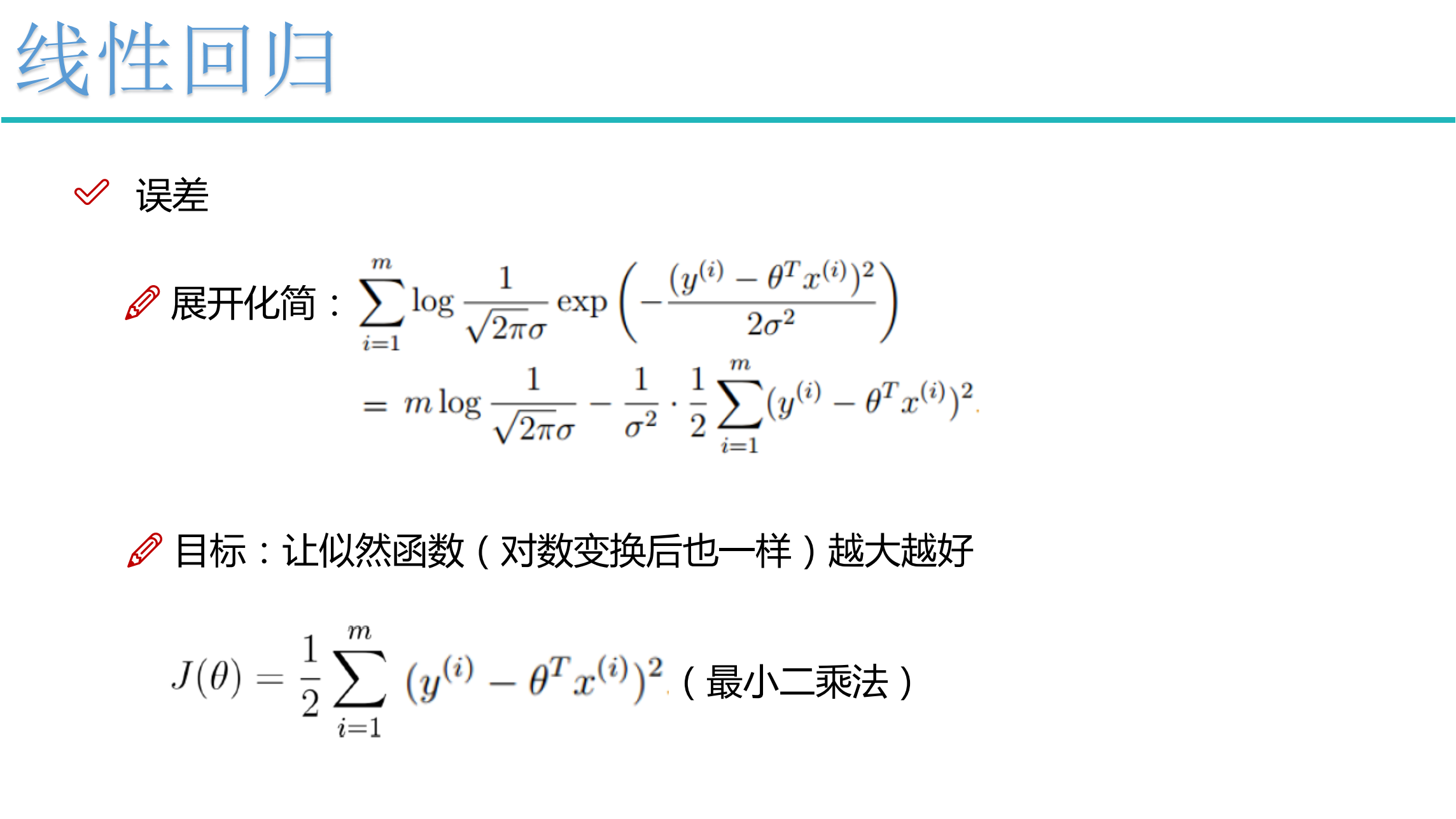
上述推导可以看出最大似然估计受到样本量的影响较大

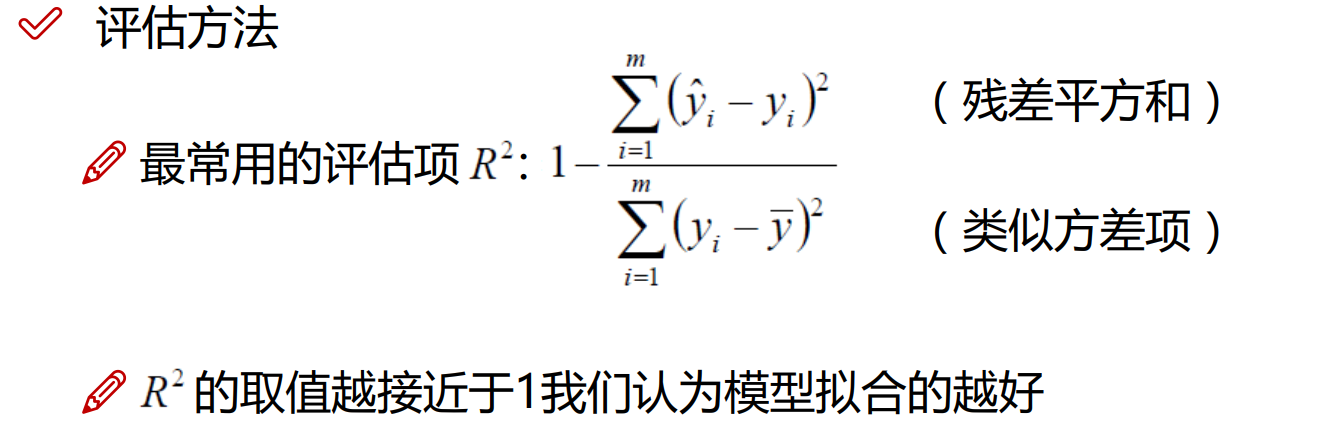
当我们知道某一组数据是服从已知分布的，那我们就可以进行估计

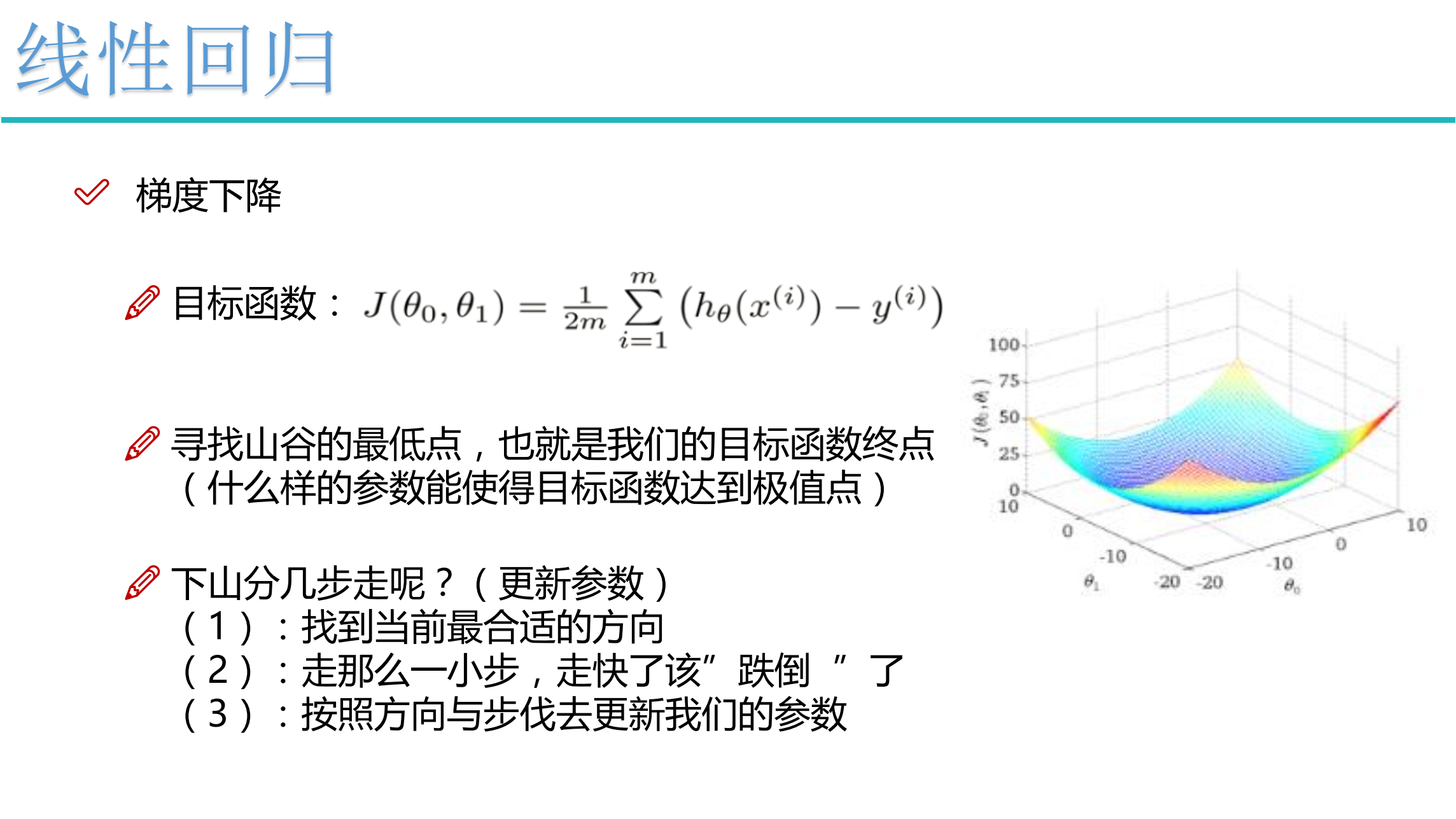
这里的似然函数本质上来理解的话

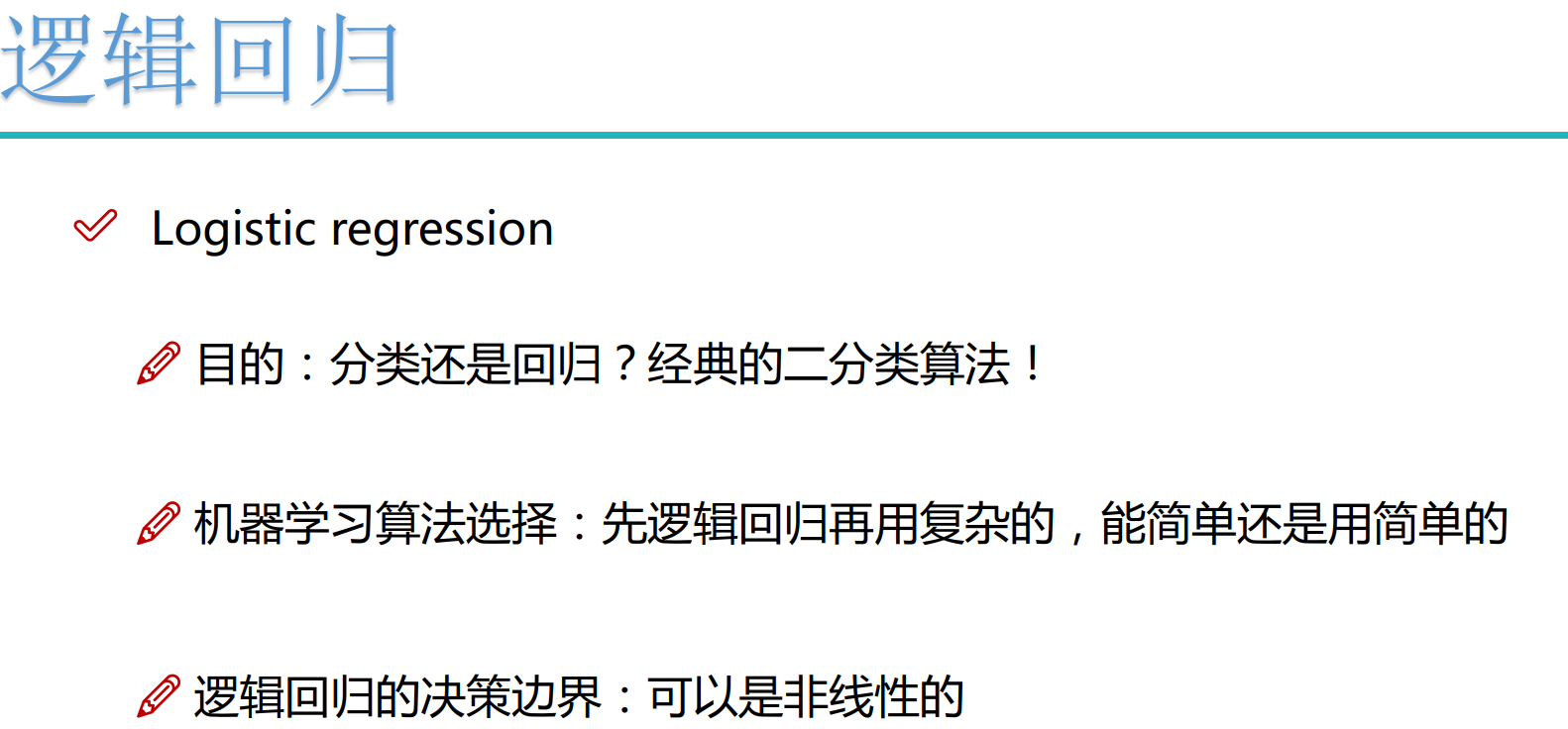
比如我们已经得到了一组预测值与真实值之间的误差值，我们知道其服从正态分布，但不知道具体参数，这里似然函数就是，不同误差值发生的概率值相乘得到的一个函数。我们要找到这个高斯分布的两个参数，这两个参数要让误差分布最接近我们真实的误差样本值，直观上理解就是，找到这样的两个参数要使得我们观测的误差样本最有可能发生就是，所有样本都发的的概率最大，就是所有样本误差发生的概率的积最大，如果每一个样本发生的概率都是0.1和每一样本发生的概率都是0.5肯定是不样的。











线性回归和逻辑回归的不同

