(提炼自周志华《机器学习》的1.4节)

设为训练集，为样本空间，为假设空间, 为目标函数, 为“算法在训练集上产生假设的概率”，为算法的“训练集外误差”，即算法在训练集之外的所有样本上的误差。

按照上述定义，有，

其中，函数为条件语句expr的真值，当expr为真时值为1，否则值为0。绝对值符号不表示绝对值，而表示0范数，也就是元素个数。

考虑二分类问题，且真实目标函数可以是任何函数。对所有可能的按均匀分布求误差之和，有

考虑到这样的函数有个，且若均匀分布，则有一半的对的预测与不一致。

要注意的是，对离散空间的函数来说，同一个函数有无数个解析式，所以不能想当然地认为，的解析式有无穷多个则有无穷多个。

因此又有

该式显示出，总误差与学习算法无关！

我们需要注意到，NFL定理有一个重要前提：所有“问题”出现的机会相同，或者所有问题同等重要。但是实际情况是，我们只关注自己正在试图解决的问题。NFL最重要的寓意是，脱离具体问题，空谈“什么学习算法、模型更好”毫无意义。因为若考虑所有潜在问题，则所有学习算法一样好。

原始论文：

Wolpert, 1996; Wolpert and Macready, 1995.