**逻辑斯蒂回归（LR），回归与最大熵模型的关系**

二项逻辑斯蒂回归公式如下：

如果我们把改写为，则

这个表达式很像的表达式。

所以，我们可以把二项逻辑斯蒂回归当做一个简单的神经网络：

图中虚线代表权重为，这样就可以保证的值恒为，这样我们对这个网络的输出层做就得到了和上面一样的表达式。由于我们可以理解为**概率**与**对数几率**成正比的映射函数，那么逻辑斯蒂回归就是将所求的值与进行对数的比较。

照这个思路，我们可以将其拓展为多项的逻辑斯蒂回归网络：

它和普通的softmax函数的区别就是存在恒等于0的几个权重。

以下所有蓝边和黄边的权重不一定相等。

概率公式为：

我们得出的结论是逻辑斯蒂回归可以看做是softmax回归的一个特例，下面讨论softmax回归和最大熵模型的关系。

最大熵模型的一个特点就是特征函数

特征函数允许人为地针对问题使用不同的特征函数，右图中对于每个不同的都会得到不同的值。提取出特征函数后，再加入权重进行调整，因此，如右图颜色一致部分所示，特征函数的权重对于每一类都是一样的。当然，我们也可以让特征函数在每类中的权值不一样，但是这样大概是没必要的，我们可以将“因类而异”的信息直接在特征函数中体现。回过头来看softmax回归，它实际上是把输入的每一维都当做一个特征，但如果它向最大熵模型一样共享权重，那么对于每一类的概率是一致的，因为左侧输入只有，不会随变化。

**极大似然法求解LR**

二项逻辑回归中，我们可以把概率简单合并到一个分布律中：

因为只能取得0和1，所以这么写可以将上面两种情况融合在一起。这样写的好处是在极大似然中求解方便。