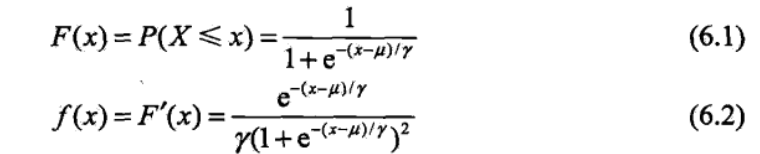
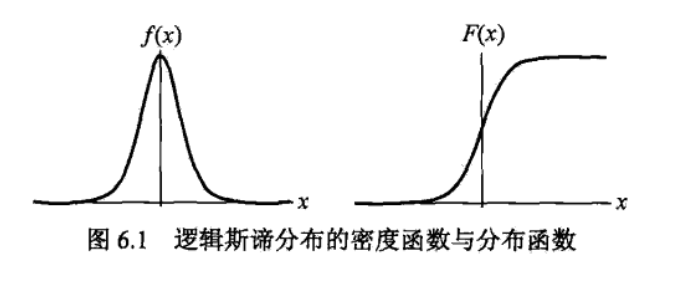
逻辑斯蒂回归是统计学习中的经典分类方法，常用于二分类，属于对数线性模型。

# 1.逻辑斯蒂分布

设X是连续随机变量，X服从逻辑斯蒂分布是指X具有下列分布函数和密度函数：



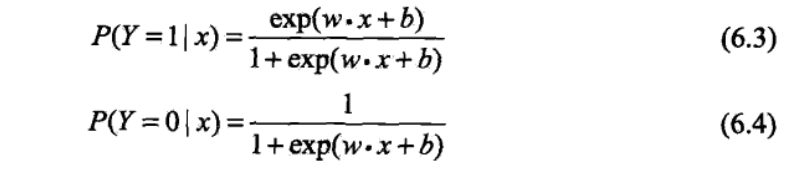
其中，μ为位置参数，λ>0为形状参数F(x)和f(x)的图形如下：



分布函数F(x)图形是一条S形曲线，以点(μ,1/2)为中心对称。

# 2.二项逻辑斯蒂回归模型

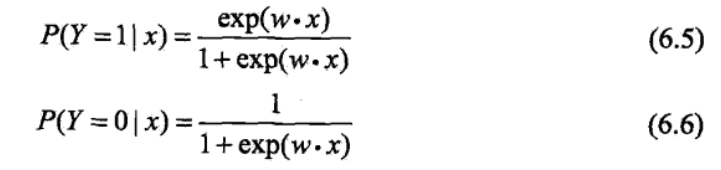
该模型是一种分类模型，由条件概率P(Y|X)表示，其中随机变量X取值为实数，Y取值为1或0.二项逻辑斯蒂回归模型的条件概率分布表示如下：



式中，w为权重向量，b为偏置向量。

对于给定的输入实例x，按照上式求得P(Y=1|X)和P(Y=0|X)，概率大的即为实例x所属的类别。

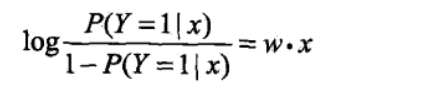
令w=(w(1),w(2),...,w(n),b)),x=(x(1),x(2),...x(n),1)，则逻辑斯蒂回归模型如下：



一个事件的几率是指该事件发生的概率和该事件不发生的概率的比值。如果事件发生概率为p,则几率为p/(1-p),事件的对数几率为:

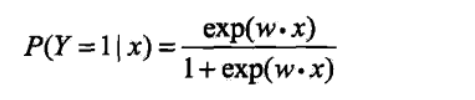
logit(p)=log(p/(1-p))

那么，根据式6.5和6.6得：



也就是说，输出y=1得对数几率是输入x的线性函数，或者说，输出y=1的对数几率是由输入x的线性函数表示的模型，即逻辑斯蒂回归模型。

那么，通过逻辑斯蒂回归模型定义6.5可以将线性函数转换为概率：

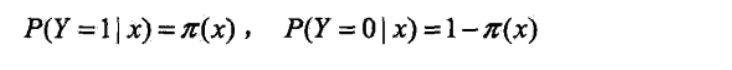


此时，线性函数的值越接近正无穷，概率值就越接近1；线性函数的值越接近负无穷，概率值越接近0.

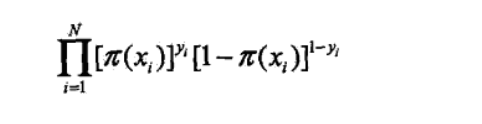
# 3.模型参数估计

对于给定训练数据集T={(x1,y1),(x2,y2),...,(xN,yN))}，xiRn,yi{0,1},，可以应用极大似然估计法估计模型参数，从而得到逻辑斯蒂回归模型。

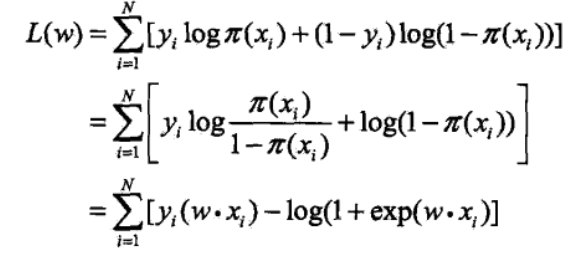
设：



似然函数为：



对数似然函数为：



对L(w)求极大值，得到w的估计值。这样问题就成了以对数似然函数为目标函数的最优化问题。逻辑斯蒂回归学习通常使用梯度下降或者拟牛顿法来更新参数：

梯度计算如下：



=

假设w的极大似然估计为，那么学到的逻辑斯蒂回归模型为:

