## 01. 损失函数与目标函数

损失函数：是定义在*单个样本上*的，算的是一个样本的误差。

代价函数：是定义在*整个训练集*上的，是所有样本误差的平均，也就是损失函数的平均。

目标函数：定义为：最终需要优化的函数。等于经验风险+结构风险（也就是代价函数+ 正 则化项）。

## 02. 损失函数分析

2分类问题：当标签为+-1的时候， fy为函数距离与置信度评估

0-1损失：最自然的想法，但是不可导:



hinge合页损失：次梯度下降法:



-1+1交叉熵：



0-1交叉熵：



**回归问题：**

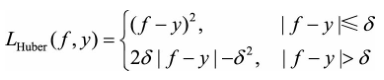
均方误差：光滑可导，对于异常值敏感，但在优化过程中更为稳定和准



绝对损失函数：对于异常值更鲁棒，但它的导数不连续使得寻找最优解的过程低效



huber：结合上面两种的优点，即是光滑的，而且对异常值没有那么敏感。



## 03 代价函数选取

一般选取对参数可微的代价函数。

回归问题一般选取MSE均方误差作为作为代价函数

分类问题一般选取交叉熵损失函数构成的代价函数。

对于logistic regression 来说，一般不适用均方误差来作为Loss Function，这是因为：均方误差损失函数一般是非凸函数（non-convex），在使用梯度下降算法的时候，容易得到*局部最优解*，而不是全局最优解。因此要选择**凸函数**（二阶导大于等于0）。