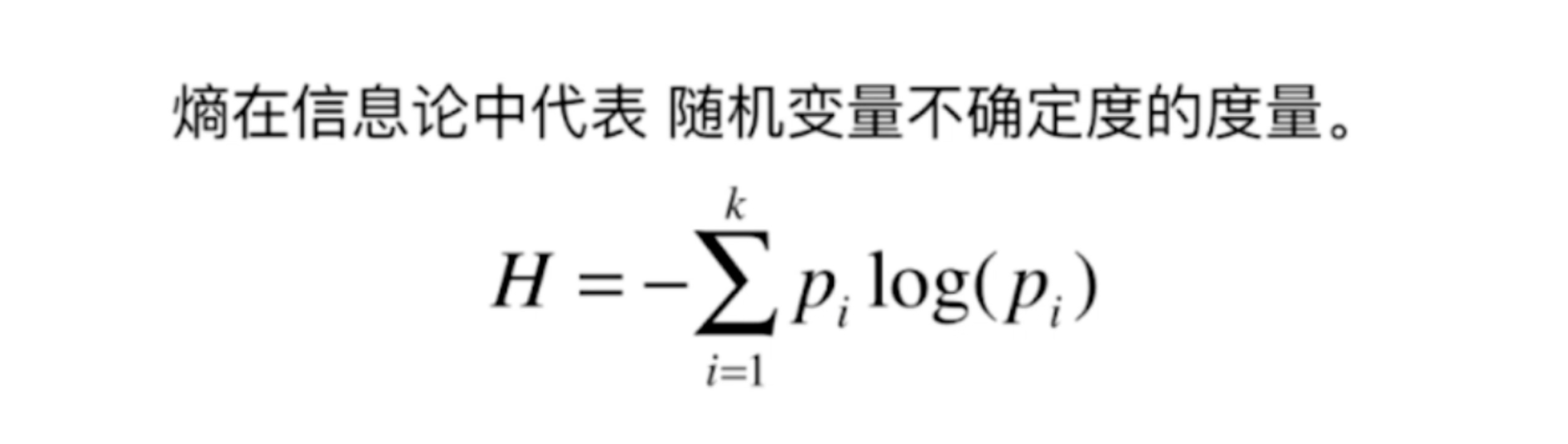


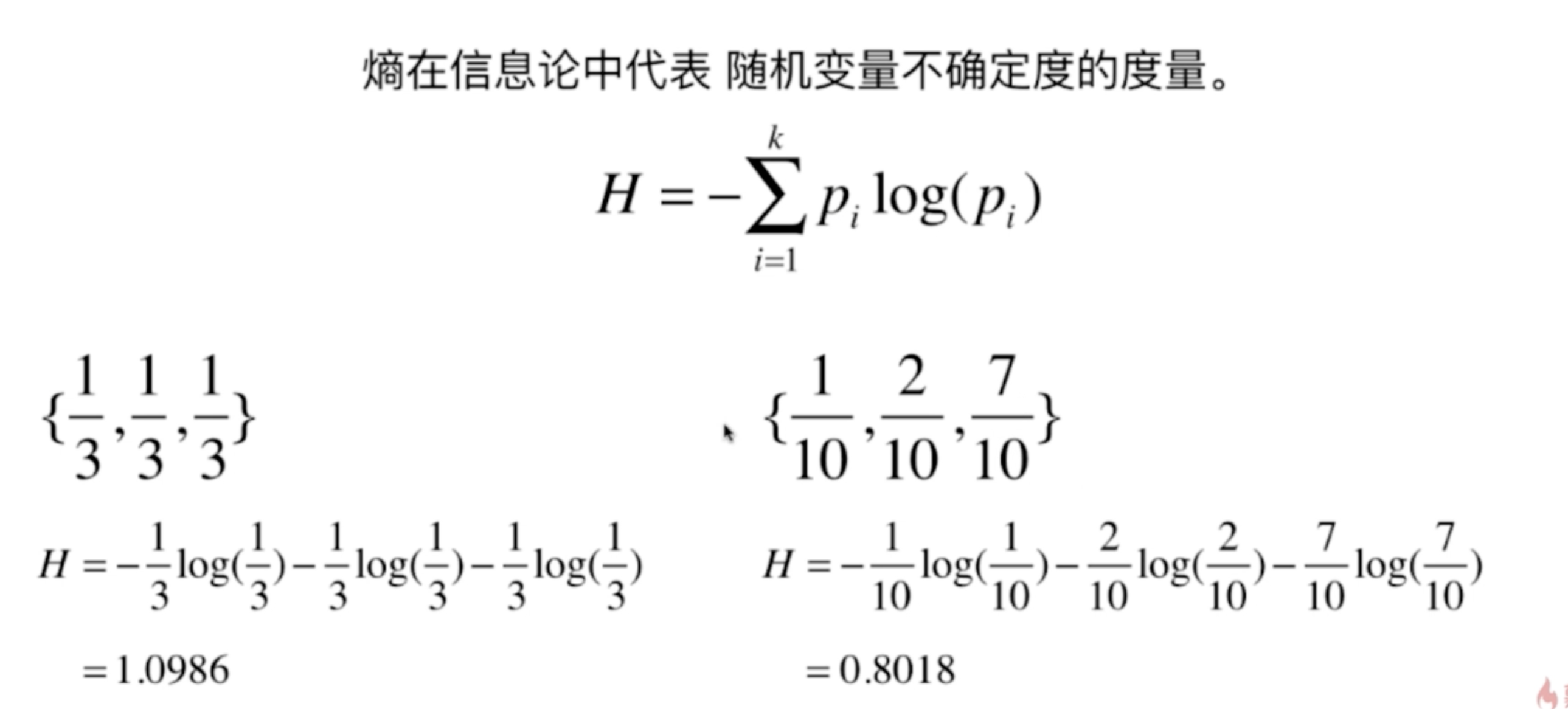
熵的概念是从物理热力学借鉴的概念。熵越大，微观粒子的无规则的运动就越剧烈，所以它们的不确定性越高。

来看看信息熵的公式：



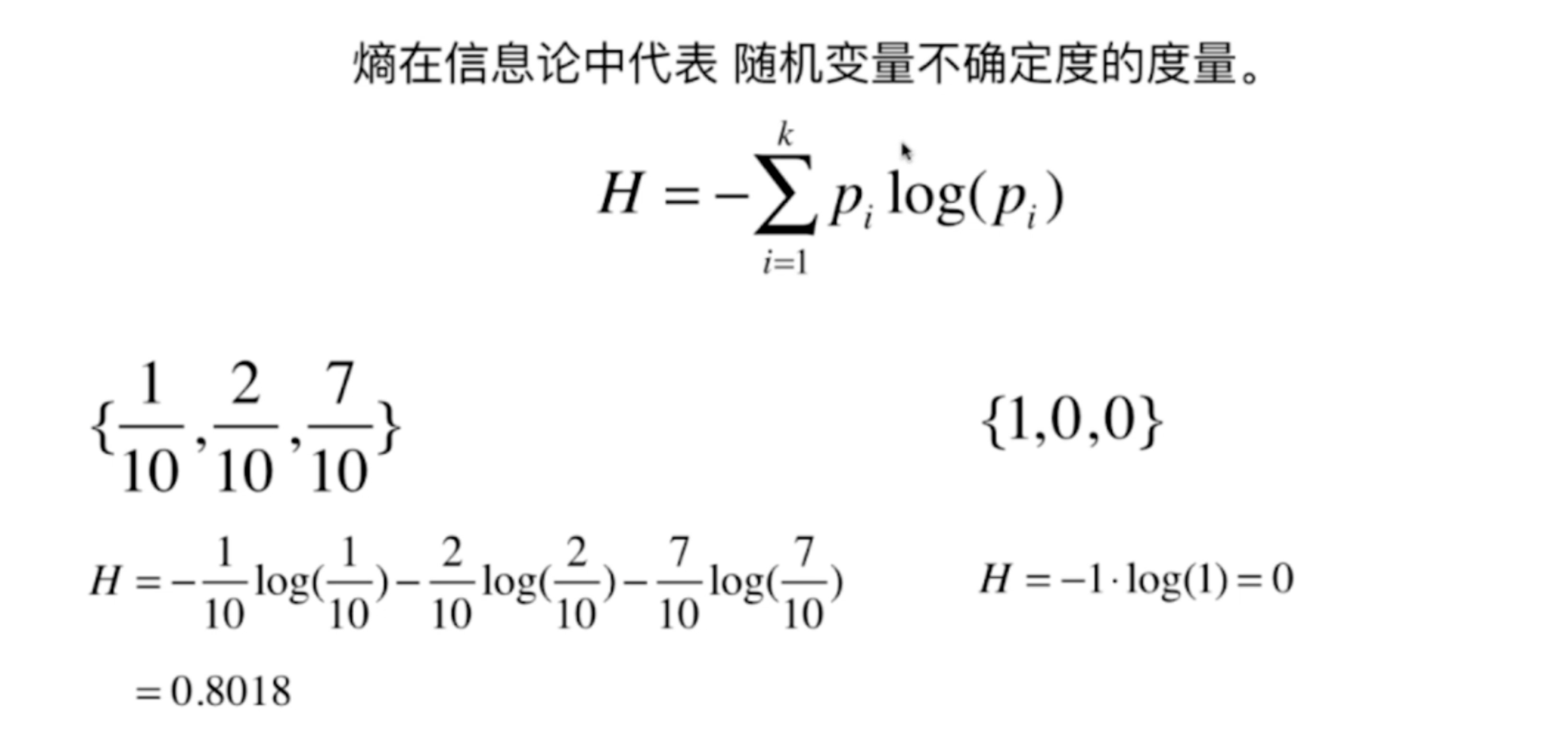
K是包含的类型的个数，p是每种类型占总样本的比例。

看个例子：



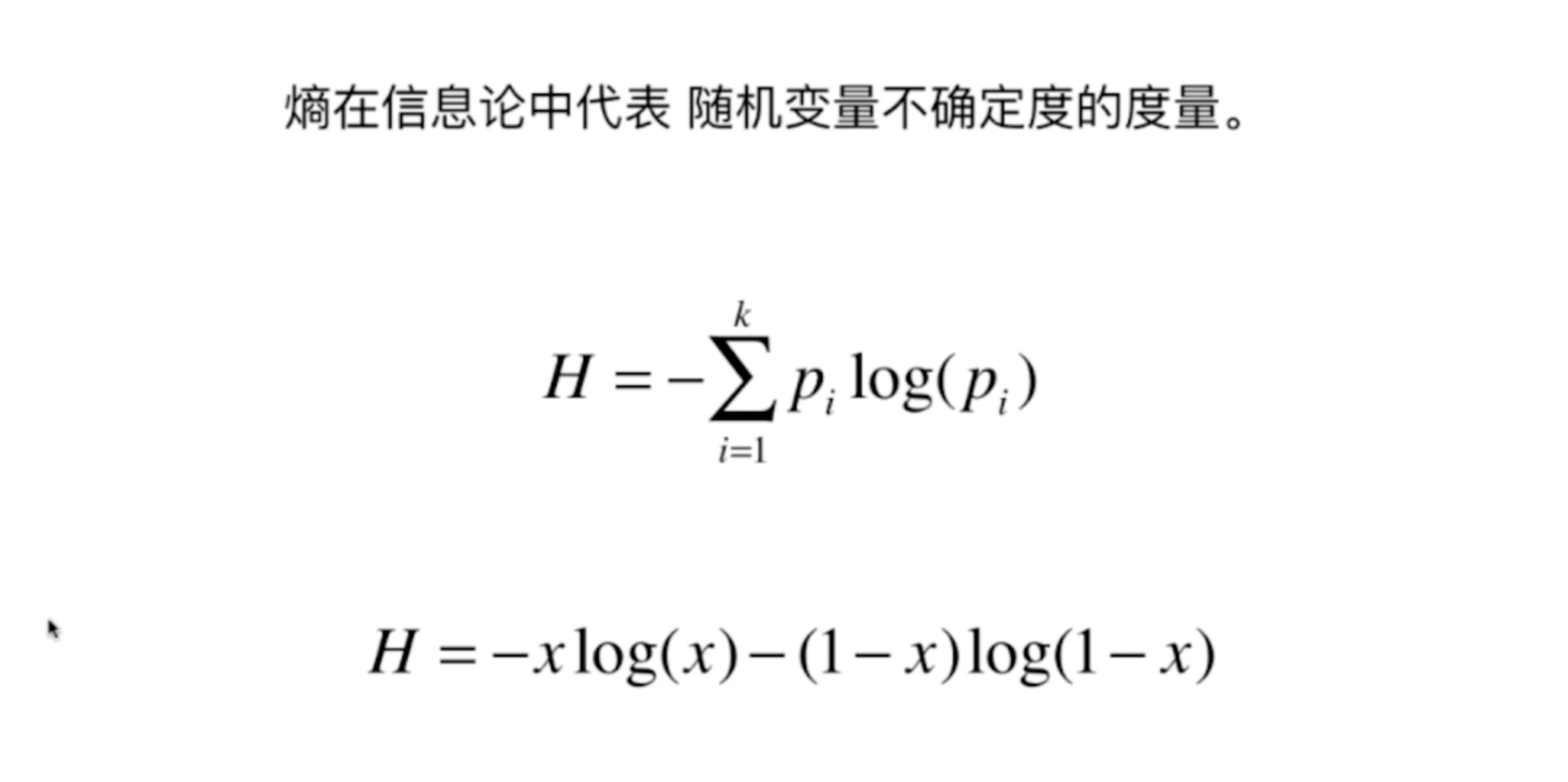
左边的信息熵大于右边，是因为右边大部分的数据都归于 7/10 更加稳定。而左边的情况是各自1/3，不确定性较高。

再来看一个极端的例子：



右边的情况是信息熵的最低值，没有任何不确定性，也就是说是最确定的。

下面用代码来看看二分类中的信息熵的公司和实验：



代码：[12-02 Entropy.ipynb]